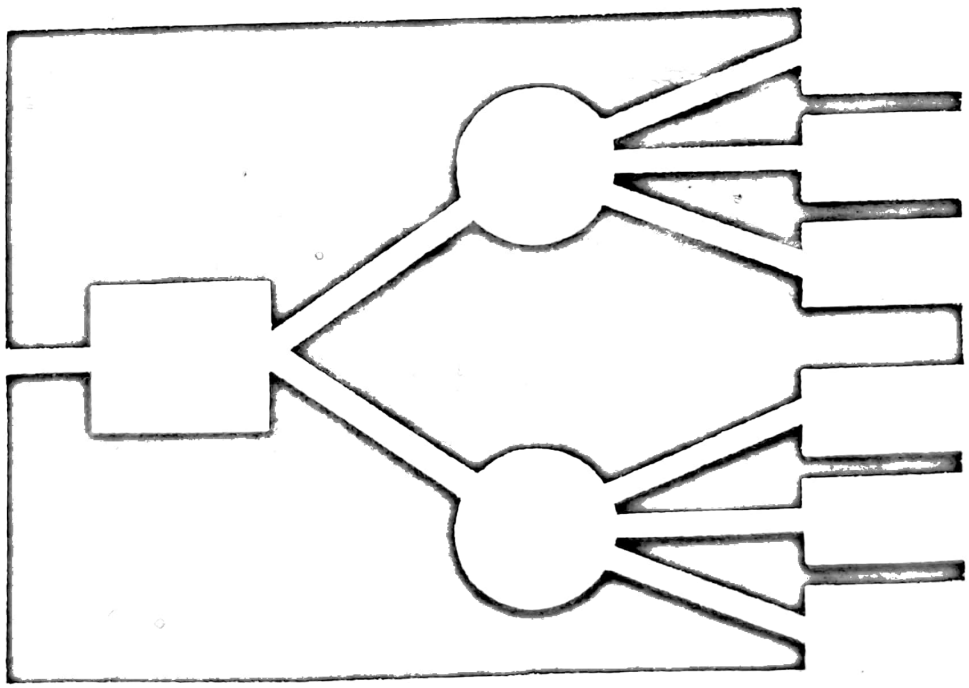
M.Malița C.Zidăroiu

INCERTITUDINE ȘI DECIZIE



Editura
științifică și enciclopedică

Scanned with OKEN Scanner

Activitatea de conducere politică, economică, militară, tehnologică sau administrativă nu poate fi concepută în afara rezolvării problemelor de luare a deciziilor. Importanța deciziei și dificultatea alegerii fac deseori necesară o analiză riguroasă.

Teoria deciziei este în esență *atitudinea științifică* față de procesul de luare a deciziilor. Părți importante ale acestui proces pot fi analizate în mod sistematic. Această analiză permite punerea în evidență a unor *legi* ale procesului de luare a deciziilor, legi ce izvorăsc din nevoia de *coerență* în abordarea diverselor aspecte ale problemei de decizie; încălcarea acestor legi poate conduce la adoptarea unor decizii ridicole sau chiar absurde.

Teoria matematică a deciziei își are izvoarele în problemele clasice ale statisticii matematice: estimație punctuală, estimație prin regiuni de încredere și testarea ipotezelor statistice. Sesizîndu-se analogiile între aceste probleme și problemele de decizie în condiții de incertitudine s-a elaborat o *teorie generală* care poate fi aplicată cu succes în cele mai diverse domenii ale activității umane.

Scanned with OKEN Scanner

MIRCEA MALIȚA ● CORNELIU ZIDĂROIU

INCERTITUDINE ȘI DECIZIE

Vol. I

(Analiza incertitudinii)

EDITURA ȘTIINȚIFICĂ ȘI ENCICLOPEDICĂ
București, 1980

Scanned with OKEN Scanner

Partea întîi

INTRODUCERE ELEMENTARĂ ÎN TEORIA DECIZIEI

Scanned with OKEN Scanner

Capitolul 1

FORMULAREA PROBLEMEI

1.1. INTRODUCERE

Viața profesională sau particulară a individului, activitatea de conducere politică, economică, militară, administrativă sau tehnologică la diverse nivele nu pot fi concepute înafara rezolvării diverselor probleme de decizie care apar în aceste domenii ale activității umane.

Unele decizii se iau fără dificultate; nu este necesar să reflectăm prea mult asupra alternativelor pentru a alege o decizie satisfăcătoare. Cele mai multe dintre deciziile pe care le ia individul în viața sa particulară aparțin acestei categorii. Alegerea programului pentru petrecerea timpului liber, a sfîrșitului de săptămînă sau a concediului de odihnă nu prezintă probleme deosebite, deși insatisfacții datorate timpului nefavorabil sau condițiilor de confort pot să apară. Evident, alegerea unei poziții profesionale sau cumpărarea unei locuințe sînt decizii mult mai importante, la care trebuie să reflectăm mai serios, ținînd seama de consecințele pe care le pot avea în viitorul mai apropiat sau mai depărtat.

Există cazuri în care importanța deciziei și dificultatea de a alege între mai multe acțiuni posibile fac necesară o analiză riguroasă. O bună parte dintre deciziile cu care sînt confruntate organismele de conducere politică, economică, militară, administrativă sau tehnologică aparțin acestei categorii. Consecințele pe care le are adoptarea unei decizii la acest nivel au o importanță atît de mare pentru o parte sau pentru toți membrii societății încît apelul la rațiune, la analiză aprofundată a diverselor alternative în scopul depistării unei acțiuni satisfăcătoare este pe deplin justificat.

Această lucrare are drept scop prezentarea unei *atitudini științifice* față de procesul de luare a deciziilor. Contribuția pe care o astfel de abordare o poate avea la procesul concret de luare a deciziilor trebuie clarificată de la început. Se afirmă

frecvent că luarea deciziilor este o activitate umană de esență politică, în care *organul de decizie* are o personalitate bine definită și că o abordare cantitativă și abstractă, bazată pe matematică, a acestei activități este inadecvată. Alexandru Macedon sau Napoleon n-au avut nevoie de o teorie a deciziei pentru adoptarea deciziilor politice și militare care i-au făcut celebri.

Sîntem de acord cu afirmația că elementul uman și politic joacă rolul esențial în adoptarea deciziilor. Totuși, există părți ale procesului de luare a deciziilor pe care putem să le analizăm în mod sistematic. Așa cum vom vedea, se poate realiza o „descompunere" a problemei de decizie în subprobleme mai simple, din a căror recombinare putem obține soluția problemei în întregime. Mai mult, această analiză ne permite să punem în evidență anumite *legi* ale procesului de luare a deciziilor, a căror încălcare poate conduce la adoptarea de decizii ridicole sau chiar absurde. Aceste legi izvorăsc din nevoia de *consistență* sau *coerență* în abordarea diverselor aspecte ale problemei de decizie. Asupra acestor aspecte vom insista peste tot în această lucrare, al cărei subiect ar putea fi descris pe scurt drept „adoptarea coerentă a deciziilor".

Un prim element al problemei de decizie îl constituie mulțimea alternativelor care pot fi adoptate de către organul de decizie. Evident, avem o problemă de decizie numai dacă există cel puțin două alternative. De obicei, sîntem interesați în adoptarea unei anumite decizii și tentați să comparăm această decizie cu alternativa contrară. De exemplu, dacă ne punem problema să mergem la cinematograf într-o anumită seară, atunci sîntem tentați să considerăm că lista deciziilor este formată din două elemente: a merge sau a nu merge la cinematograf. De cele mai multe ori, acest mod de a constitui lista de alternative este inadecvat. În exemplul de mai sus, dacă nu mergem la cinematograf, putem considera cîteva alternative: citim o carte, urmărim programul la televizor sau facem o plimbare în parc. Lista deciziilor astfel constituită reprezintă un evantai de posibilități mai bine structurat, din care alegerea se poate face mai ușor.

Observații similare pot fi făcute relativ la problemele de alocare a resurselor sau de investiții, unde structurarea alternativelor la o anumită decizie considerată inițial are o importanță deosebită. Problema constă nu în a spune „da" sau „nu" la

o anumită propunere de investiție, ci de a examina și de a evalua cu mare atenție alternativele la această propunere.

Elaborarea listei de alternative disponibile este o sarcină dificilă și dintr-un alt punct de vedere. Această listă trebuie să fie *completă*, alegerea unei anumite decizii putîndu-se face numai din mulțimea alternativelor care au fost incluse în această listă. Depistarea tuturor alternativelor care prezintă interes este o operație plină de dificultăți, care presupune o bună cunoaștere a domeniului, dar și abilitatea de a avea idei noi; multe succese în luarea deciziilor se datorează mai mult ultimului factor menționat mai sus și mai puțin unei alegeri adecvate dintr-o mulțime dată de decizii. O listă (mulțime) de decizii completă va fi numită și *exhaustivă*.

Nevoia de a gîndi la o decizie anumită prin prisma alternativelor disponibile și necesitatea de a depista toate deciziile care sînt relevante pentru problema dată constituie fără îndoială pași înainte față de abordarea tradițională a problemei de decizie.

Din motive de simplitate a analizei, vom considera în această parte introductivă că din lista exhaustivă de decizii alegem una și numai una. Cu alte cuvinte, alegerea unei anumite alternative exclude alegerea oricărei alte decizii de pe listă. Deciziile dintr-o astfel de listă vor fi numite *exclusive*.

Există probleme de decizie în care lista completă a alternativelor nu are această ultimă proprietate. Să considerăm, de exemplu, problema selectării unui meniu dintr-o listă de bucate a unui restaurant, care conține o listă de gustări, o listă pentru primul fel, o listă pentru al doilea fel, o listă de deserturi și o listă de băuturi. Evident, lista de bucate nu conține alternative exclusive, deoarece alegerea unei gustări nu exclude alegerea celorlalte feluri ce vor fi consumate. Este clar că putem concepe și în acest caz o listă de decizii exclusive, dacă se consideră toate meniurile posibile formate din această carte de bucate. Vom vedea mai tîrziu că astfel de *probleme de alegere secvențială* pot fi descompuse în probleme mai simple, fiecare avînd o listă completă de decizii exclusive. În exemplul de mai sus, problema alegerii unui meniu se descompune în problema alegerii unei gustări, apoi a unui prim fel, apoi a unui al doilea fel ș.a.m.d. Este totuși indicat să începem cu o singură problemă de decizie.

Scanned with OKEN Scanner

În concluzie, vom considera o listă de decizii *exhaustivă* și *exclusivă* care conține un număr m de alternative, pe care le notăm prin $d_1, d_2, \ldots, d_m$, ordinea de numerotare a lor fiind arbitrară. Mulțimea de decizii este deci $D = \{d_1, d_2, \ldots, d_m\}$.

În cele mai multe cazuri mulțimea D a deciziilor este finită, așa cum am considerat mai sus, dar există și situații în care D nu este finită; exemple din această categorie de probleme vor apărea în cursul acestei lucrări.

Problema de decizie constă în selectarea unei alternative $d^* \in D$ care să fie, într-un anumit sens, „cea mai bună". În cazul cînd nu există incertitudine asupra consecințelor care decurg din adoptarea unei alternative sau a alteia, selecția celei mai bune decizii nu comportă în general probleme. De exemplu, dacă prietenul în al cărui gust avem încredere deplină ne recomandă călduros un film, vom merge la cinematograf; acasă, cartea pe care o citim nu este prea interesantă, programul la televizor din această seară este de asemenea neinteresant, iar timpul nu este potrivit pentru o plimbare în parc.

Sînt desigur situații în care, deși nu există nici un fel de incertitudine, selectarea „celei mai bune" decizii este o operație extrem de dificilă, dacă nu chiar imposibilă. Pe de o parte, aceasta se poate datora unor *dificultăți tehnice*, legate de volumul și complicația calculelor necesare pentru depistarea celei mai bune alternative. De exemplu, în jocul de șah se poate constitui, în principiu, o listă a tuturor deciziilor, formate din toate mutările posibile permise în fiecare fază a jocului, în funcție de mutările posibile ale celuilalt jucător. Fiecare astfel de decizie conduce la cîștig, remiză sau pierdere și problema este să se selecteze o decizie care să conducă la cîștig (dacă o astfel de decizie există!). Motivul pentru care această selecție nu este posibilă este volumul calculelor necesare în acest scop, care depășește cu mult posibilitățile celor mai moderne calculatoare. Evident, pentru jocuri mai simple, depistarea celor mai bune mutări se poate face mai simplu; există jocuri care nu mai prezintă interes decît pentru începători, decizia cea mai bună putînd fi calculată ușor cu posibilitățile actuale.

În lucrarea de față nu vom considera aceste dificultăți tehnice. Există numeroase domenii în care dificultățile tehnice au fost depășite cu succes; programarea liniară și neliniară,

Scanned with OKEN Scanner

programarea dinamică sînt cîteva dintre instrumentele cu ajutorul cărora s-au surmontat dificultăți tehnice remarcabile*.

Pe de altă parte, dificultatea alegerii poate să fie generată de faptul că diversele consecințe pe care le are adoptarea unei alternative sau a alteia nu sînt comparabile din punctul de vedere al organului de decizie. De exemplu, pentru cumpărarea unei locuințe avem următoarele alternative: (d_1) cumpărarea unui apartament cu trei camere confort sporit în cartierul Titan, (d_2) cumpărarea unui apartament cu patru camere confort I în cartierul Berceni și (d_3) cumpărarea unui apartament de trei camere confort II în cartierul Drumul Taberei. Din punctul de vedere al confortului, preferăm d_1 lui d_2, iar d_2 lui d_3. Din punctul de vedere al suprafeței locative, preferăm d_2 lui d_1 și d_1 lui d_3. În sfîrșit, din punctul de vedere al apropierii de locul de muncă, preferăm d_3 lui d_2 și d_2 lui d_1. Ce alegere facem? Dificultăți de această natură sau similare acestora apar deseori în procesul de luare a deciziilor; partea a treia a acestei lucrări este consacrată în întregime rezolvării problemelor de acest tip.

În sfîrșit, trebuie să subliniem că elementul care complică foarte mult problemele de decizie este *incertitudinea;* cele mai multe probleme de decizie comportă raționamente asupra unor fenomene și situații incerte. În cele ce urmează, spre deosebire de accepțiunea curentă a termenului, vom utiliza noțiunea de *eveniment* pentru a desemna apariția sau producerea (neproducerea) unui anumit fenomen sau situație. Un eveniment este numit *sigur* sau *cert* dacă sîntem informați că s-a produs sau se va produce în viitor cu siguranță; în caz contrar vorbim despre *evenimente incerte*. Cu alte cuvinte, faptul că un eveniment este cert sau incert este *o apreciere* a organului de decizie, pe baza informației disponibile la un moment dat, și nu este neapărat o caracteristică *intrinsecă* sau *obiectivă* a acestui eveniment. De exemplu, dacă nu avem nici o informație despre filmul pe care dorim să-l vedem, rezultă că el poate fi un film bun (evenimentul incert ω_1), un film mediocru (evenimentul ω_2) sau un film slab (evenimentul ω_3). Ca în foarte multe alte probleme de decizie, în acest exemplu incertitudinea își are izvorul

---

* Pentru o introducere în acest domeniu, vezi M. MALIȚA, C. ZIDĂROIU (1975).

în lipsa sau insuficiența informației disponibile în momentul în care are loc alegerea alternativei.

Pe de altă parte, există numeroase exemple de evenimente incerte „în mod obiectiv". De exemplu, nimeni nu poate să spună dinainte dacă rezultatul la aruncarea unei monezi va fi stema sau valoarea. În acest caz incertitudinea este datorată *întîmplării;* apariția stemei (evenimentul incert ω_1) sau apariția valorii (evenimentul incert ω_2) sînt numite din aceste motive și *evenimente întîmplătoare.*

Indiferent de originea incertitudinii, în orice problemă de decizie vom construi o listă a tuturor evenimentelor incerte care pot afecta alegerea unei decizii. Ca și în cazul deciziilor, această listă trebuie să fie *completă (exhaustivă),* iar evenimentele incerte trebuie să fie *exclusive;* cu alte cuvinte, unul și numai unul dintre aceste evenimente s-a produs sau se va produce cu siguranță. Vom nota prin $\omega_1, \omega_2, \ldots, \omega_n$ evenimentele incerte din această listă și prin $E = \{\omega_1, \omega_2, \ldots, \omega_n\}$ mulțimea acestor evenimente incerte. În cele mai multe cazuri, E este o mulțime finită, dar există și situații în care E este infinită, așa cum vom avea ocazia să constatăm în exemplele ce vor fi date în continuare.

Construirea unei liste exhaustive și exclusive de evenimente incerte este în multe cazuri o operație dificilă și delicată, care presupune o bună cunoaștere a factorilor care ar putea afecta selectarea unei decizii sau a alteia. Există numeroase exemple, în special din domeniul militar, care demonstrează consecințele catastrofale pe care le poate avea asupra unei decizii ignorarea unei singure eventualități. Din acest motiv, principala preocupare este ca lista evenimentelor incerte să fie *exhaustivă.* Aceasta presupune din partea organului de decizie abilitatea de a depista toți factorii care pot influența alegerea unei decizii sau a alteia. Lista conținînd toate evenimentele incerte relevante poate fi apoi transformată într-o listă de evenimente exclusive utilizînd procedee similare celor folosite în același scop pentru obținerea listei de decizii.

Studiul incertitudinii a condus la elaborarea unei teorii matematice, inițiată în urmă cu trei secole, și cunoscută sub numele de *teoria probabilităților.* Partea a doua a lucrării noastre este dedicată în întregime prezentării principalelor elemente ale acestei teorii matematice, elemente relevante pentru rezolvarea problemelor de decizie.

Scanned with OKEN Scanner

În concluzie, o problemă de decizie constă în esență în selectarea unei decizii din mulțimea $D = \{d_1, d_2, ..., d_m\}$, fără să știm care dintre evenimentele incerte din mulțimea $E = \{\omega_1, \omega_2, ..., \omega_n\}$ a avut sau va avea loc. Desigur, în această selecție un rol important joacă, pe de o parte, preferințele pe care le avem față de consecințele care pot să apară în urma adoptării unei alternative sau a alteia. Pe de altă parte, informația disponibilă în ceea ce privește evenimentele incerte, precum și informația care ar putea fi obținută în plus, contribuie de asemenea, într-o măsură importantă, la rezolvarea problemei de decizie.

1.2. EXEMPLE

Înainte de a trece la formularea detaliată a problemei de decizie, vom prezenta o serie de exemple concrete de astfel de probleme. Pe de o parte, ilustrăm în acest mod gama impresionantă de domenii ale activității umane în care ideile generale expuse mai sus se materializează. Pe de altă parte, aceste exemple vor îmbogăți ideile generale, conducîndu-ne la surprinderea structurii comune a acestor probleme și, implicit, la posibilitatea formulării adecvate a modelului general al problemei de decizie.

1.2.1. PROBLEME DE CERCETARE ȘI DEZVOLTARE

Una dintre problemele majore cu care sînt confruntate organismele de conducere economică la diverse niveluri este problema selectării temelor de cercetare, a proiectelor pentru produse sau tehnologii noi, a momentului lansării unui produs, a măririi capacităților de producție destinate acestui nou produs, a cooperării cu alte întreprinderi în realizarea acestui nou produs ș.a.m.d. Vom da în continuare cîteva exemple din această categorie importantă de probleme de decizie.

Scanned with OKEN Scanner

(a) *Selectarea unui program de cercetare în domeniul energetic*

Un organism de conducere economică însărcinat cu probleme de dezvoltare a resurselor energetice urmează să analizeze un proiect de cercetare privind dezvoltarea reactorilor nucleari cu regenerare rapidă, care urmează să se desfășoare pe o perioadă de zece ani și pentru care sînt necesare investiții importante. Caracteristica principală a acestor reactori nucleari constă în faptul că sînt o sursă de combustibil nuclear pentru centralele electrice nucleare clasice, consumînd mai puțin combustibil nuclear natural decît cantitatea pe care o produc și deci reducînd pe această cale dependența de resursele naturale. Utilizarea acestor reactori ar permite producerea economică a electricității, chiar în cazul în care resursele s-ar găsi din ce în ce mai greu; dealtfel, rezervele naturale de combustibil nuclear cunoscute în prezent ar urma să se epuizeze într-o perioadă de opt ani, dacă cererea de energie nucleară crește conform previziunilor actuale.

Pe de altă parte, sînt încă multe probleme tehnice nerezolvate în domeniul reactorilor cu regenerare rapidă. Din acest motiv există incertitudini serioase în ceea ce privește construirea în decurs de zece ani a unui reactor nuclear cu regenerare rapidă care să fie atît economic cît și sigur în funcționare.

Renunțarea la acest proiect ar avea drept consecință creșterea rapidă a costului electricității produse în centralele nucleare clasice, datorată în special creșterii previzibile a costului combustibilului nuclear natural.

Dacă se decide lansarea acestui proiect, urmează să se decidă și mărimea investiției. Se poate concepe realizarea unui proiect de mare anvergură, pe o perioadă mai scurtă, cu posibilitatea renunțării la cercetări în momentul descoperirii de noi rezerve naturale de combustibil nuclear. Este de asemenea posibil să concepem realizarea unui proiect de dimensiuni mai modeste, dar pe o perioadă mai lungă, așteptînd descoperirea de noi resurse naturale de combustibil.

Odată adoptată una sau alta dintre aceste variante, sînt încă alte alternative între care trebuie să se decidă. Există mai multe descoperiri tehnice importante, privind în special conductibilitatea calorică a unor materiale, care ar putea sta la baza construirii unui reactor nuclear cu regenerare rapidă.

Scanned with OKEN Scanner

Trebuie să se decidă deci dacă fondurile de investiții trebuie concentrate pe una dintre aceste idei, sau trebuie repartizate pe proiecte paralele, vizînd cercetări noi în legătură cu toate sau, eventual, numai cu o parte dintre descoperirile menționate.

(b) *Lansarea unui nou produs*

Un material plastic nou a fost descoperit de laboratorul de cercetare al unei mari întreprinderi chimice. Fiind încă incomplet pus la punct, intrarea în fabricație imediată este riscantă. Pe de altă parte, este de așteptat ca pe piața externă să apară produse similare, intrarea devreme în producție avînd deci un rol foarte important pentru desfacerile pe piața externă în viitor.

Incertitudinile în această problemă sînt numeroase: volumul vînzărilor în viitor, comportarea prețului de vînzare pe piața externă, volumul de vînzări pe piața externă al întreprinderii noastre pentru o anumită politică asupra prețului și pentru un anumit moment al intrării pe piața externă etc.

Întreprinderea trebuie să decidă dacă și cînd va lansa materialul plastic în producție. În caz afirmativ, va trebui să decidă dacă va coopera în producție cu parteneri interni sau externi sau dacă și cînd va fi necesară o nouă unitate de producție pentru acest material plastic. În caz negativ, va trebui să evalueze avantajele pe care le-ar obține din valorificarea unor patente sau *know-how* privind acest material plastic.

(c) *Reproiectarea unui produs*

Laboratorul de cercetare al unei întreprinderi producătoare de automobile a reușit reproiectarea pompei de benzină, realizînd astfel o reducere a costului acesteia. Modificările cerute în procesul tehnologic de introducerea în fabricație de serie a pompei noi, împreună cu costul reproiectării acesteia, conduc la un cost total de modificare care este cunoscut.

Performanța pompei vechi depinde de durata sa de funcționare; probabilitățile de defectare a acesteia după 1, 2, 3, ... ani de funcționare sînt bine cunoscute din experiența acumu-

Scanned with OKEN Scanner

lată pînă în prezent. Sînt de asemenea cunoscute costurile implicate de înlocuirea acesteia după 1, 2, 3, ... ani de funcționare.

Pe de altă parte, informația privind noua pompă de benzină este mult mai săracă. Se știe numai că defectarea acesteia poate avea zece cauze diferite, fiecare dintre ele avînd anumite șanse (cunoscute) de a contribui la defectare. Corpul tehnic poate furniza de asemenea probabilitățile de defectare după 1, 2, 3, ... ani de funcționare a noii pompe pentru fiecare dintre aceste cauze de defectare.

Este cunoscută de asemenea producția de automobile în următorii zece ani, sub forma unei prognoze.

Alternativele care ar putea fi luate în considerare sînt următoarele: (d_1) continuarea folosirii pompei vechi și (d_2) utilizarea pompei noi la producția de automobile din anul următor și apoi, pe baza informației obținute despre performanțele pompei noi, continuarea folosirii acesteia sau revenirea la cea veche.

1.2.2. PROBLEME DE NEGOCIERE ECONOMICĂ

(a) *O problemă de alegere a furnizorului*

O întreprindere economică are de ales între doi furnizori A și B pentru achiziționarea unei piese necesară la realizarea unuia dintre produsele sale. O piesă este considerată bună dacă două dintre dimensiunile sale, notate prin x și y, satisfac condițiile $x' \leqslant x \leqslant x''$ și $y' \leqslant y \leqslant y''$, unde x', x'', y' și y'' sînt limite de toleranță specificate. Dacă $x < x'$ sau $y < y'$, atunci piesa este rebut și se aruncă. Dacă $x > x''$ sau $y > y''$, atunci piesa trebuie reprelucrată, ceea ce presupune un cost unitar c sau $2c$, după cum numai una sau ambele dimensiuni ale piesei trebuie rectificate.

Cei doi furnizori oferă piesele la costurile unitare $c_A < c_B$, dar există incertitudini serioase privind calitatea pieselor furnizate de A, cu care s-a colaborat mai puțin. În ceea ce privește furnizorul B, care este mai bine cunoscut, se știe că producția

sa de piese se structurează în mai multe calități, pentru care există cîte o repartiție cunoscută a defectelor.

Problema care se pune este să alegem între cei doi furnizori, făcînd eventual investigații suplimentare privind calitatea pieselor furnizate de A.

(b) *O problemă de acceptare a unui contract*

Unei întreprinderi producătoare i se oferă un contract privind construirea unui sistem complex format din n elemente și menținerea lui în stare de funcționare pe o perioadă de garanție de m ani. Orice avarie a acestui sistem care apare în perioada de garanție trebuie remediată de producător, care va înlocui imediat elementul defect, operație care comportă un preț unitar c_0 cunoscut. Instalarea sistemului are un preț de cost k_0, iar valoarea contractului oferit este c.

Incertitudinea este generată în acest caz de necunoașterea ratei de defectare λ a unui element, deși producătorul știe că procesul de defectare urmează o lege Poisson * de parametru λ. Personalul tehnic al întreprinderii poate furniza o densitate de probabilitate inițială pentru λ, iar experimente suplimentare privind duratele de funcționare ale elementelor sistemului ar putea aduce date suplimentare în această direcție.

Problema care se pune este dacă acest contract se acceptă sau nu de întreprinderea producătoare, iar în caz afirmativ dacă se fac sau nu experimentări suplimentare și cîte anume.

(c) *O problemă de licitație*

O întreprindere specializată în construcții industriale licitează împotriva altor competitori pentru obținerea unui contract privind construirea într-o altă țară a unei fabrici de materiale plastice.

Costul c al construirii acestei uzine de către întreprinderea noastră este incert, ținînd seama de o serie de aspecte necunoscute, cum ar fi, de exemplu, costul forței de muncă autoh-

* Vezi § 12.7.

tone, costul unor materiale de construcție, comportarea prețurilor pe piața mondială pe durata realizării construcției etc.

Pe de altă parte, cea mai mică ofertă k_0 pe care o pot face ceilalți competitori pentru obținerea acestui contract este de asemenea incertă, din motive evidente.

Problema constă în determinarea ofertei k_1 pe care urmează să o facă întreprinderea noastră în vederea obținerii acestui contract. Este clar că, dacă $k_1 > k_0$, contractul se pierde. Pe de altă parte, dacă oferta k_1 este prea mică, atunci adjudecarea contractului s-ar putea dovedi nerentabilă, putînd duce chiar la pierderi pentru întreprinderea noastră.

(d) *O problemă de achiziție de utilaj*

O întreprindere de produse electronice intenționează să achiziționeze de pe piața externă un utilaj necesar dezvoltării capacităților sale de producție. Prețul curent pe piața externă al acestui utilaj este c_0.

Unul dintre furnizorii externi oferă utilajul la prețul $c_1 < c_0$. Problema care se pune este dacă întreprinderea noastră cumpără sau nu acest utilaj.

Furnizorul extern care oferă acest utilaj este mai puțin cunoscut, calitatea produselor sale de acest gen nefiind încă bine precizată. Se știe totuși că 90% dintre produsele sale se pot considera fără defecte, iar 10% au defecte, a căror remediere comportă un preț k. Se poate de asemenea negocia o clauză de garanție, care asigură remedierea defectelor care ar putea să apară, dar care presupune un preț de achiziție c_1', mai mare decît cel prevăzut inițial.

(e) *Controlul calității loturilor de produse*

Pentru acceptarea de către beneficiar a unui lot de produse, acestea trebuie să satisfacă anumite cerințe de calitate. În unele cazuri, controlul fiecăruia dintre aceste produse este posibil și puțin costisitor. În astfel de situații se poate opta pentru controlul întregului lot și apoi se poate accepta sau respinge acest lot pe baza cunoașterii proporției exacte $p \in [0, 1]$ de defecte din lot.

Scanned with OKEN Scanner

În cele mai multe cazuri însă, acest procedeu este inaplicabil din diverse motive (timpul necesar, costul controlului, control distructiv etc.) și decizia de acceptare sau respingere a lotului se ia pe baza concluziilor obținute din examinarea unui *eșantion*, luat la întîmplare din acest lot. Acest tip de control este numit și *controlul statistic al calității**.

Este clar că incertitudinea își are în acest caz sursa în necunoașterea adevăratei proporții $p \in [0, 1]$ de defecte din lot.

Informații despre p pot fi obținute cu ajutorul eșantionului prelevat din lot, dar și pe baza experienței trecute privind loturile similare livrate de același furnizor. Este clar că beneficiarul trebuie să decidă dacă și cîte produse din acest lot urmează să fie testate și apoi dacă lotul se acceptă sau nu. Fiecare dintre aceste decizii are diverse consecințe, care depind evident de adevărata proporție p a produselor defecte din lot. Este uzual ca preferințele între aceste consecințe să fie exprimate prin pierderile bănești generate de fiecare decizie în parte, ca funcții de proporția p a defectelor din lot.

1.2.3. PROBLEME MEDICALE

Precizarea diagnosticului medical, stabilirea unui tratament medical adecvat sau stabilirea calităților terapeutice ale medicamentelor sînt cele mai importante probleme de decizie care apar în domeniul medical. Vom prezenta în cele ce urmează un singur exemplu din acest domeniu, și anume tratamentul unei infecții pulmonare.

Medicul nu cunoaște dacă boala pacientului său este cauzată de un virus sau de un streptococ. În cazul unei infecții virale, tratamentul constă în repaus, vitamine, aspirină etc. În cazul streptococului se impune un tratament cu antibiotice.

Consecințele aplicării inadecvate a tratamentului pot fi serioase. Aplicarea antibioticelor în cazul infecției virale poate conduce la o indispoziție serioasă pe timp de aproximativ o săptămînă, iar în cazuri rare chiar la moartea pacientului.

---

* Pentru detalii, vezi M. MALIȚA, C. ZIDĂROIU, *op. cit.*

Scanned with OKEN Scanner

Pe de altă parte, dacă nu se aplică antibioticul în cazul infecției streptococice, pot să apară complicații grave, cum ar fi nefrita și reumatismul.

Se pot obține informații suplimentare privind prezența streptococului cu ajutorul unui test bazat pe o cultură a florei bacteriene prelevată de la pacient. Testul este însă scump, necesită un timp relativ îndelungat și, în plus, este imperfect, deoarece bacteriile pot muri înainte de a fi transplantate în mediul de cultură, iar prezența streptococului în cultură nu demonstrează că el este în mod necesar cauza infecției.

Medicul are la dispoziție următoarele alternative: (d_1) prescrierea antibioticelor, (d_2) tratarea infecției ca infecție virală, (d_3) utilizarea testului descris mai sus și tratarea corespunzătoare a bolii, în urma aflării rezultatului acestui test, și (d_4) administrarea antibioticului și efectuarea în paralel a testului, urmînd ca tratamentul să fie întrerupt în cazul rezultatului negativ al testului.

Care dintre aceste decizii va fi adoptată de medic?

1.3. STRUCTURA COMUNĂ

Cele cîteva probleme de decizie prezentate mai sus pot contribui la formarea unei idei generale privind structura comună a problemelor de decizie. În fiecare dintre aceste probleme există un *organ de decizie* care trebuie să facă o alegere sau o secvență de alegeri dintr-o *mulțime dată de alternative*. *Consecințele* fiecărei decizii depind de un *eveniment incert* sau de o *stare a naturii*. Organul de decizie are la dispoziție *informații* privind incertitudinile din problema sa. Mai mult, pot fi obținute informații suplimentare în acest sens pe baza efectuării unor *experimente**.

Teoria deciziei care va fi prezentată în această lucrare pornește de la ipoteza că organul de decizie *dorește* să aleagă

* Vom utiliza noțiunea de experiment în cea mai largă accepțiune a sa: vom înțelege deci prin experiment o acțiune sau operație efectuată în scopul descoperirii unui anumit fapt. Gama operațiilor din această categorie conține toate observațiile privind fenomenele naturale, tehnologice, sociale, politice etc.

Scanned with OKEN Scanner

o alternativă pentru efectuarea experimentelor pe care le consideră necesare și pentru luarea unei decizii care să fie *coerentă* cu preferințele sale privind consecințele care pot să apară și cu opiniile sale privind evenimentele incerte.

Surprinderea structurii comune a problemelor de decizie ne permite să formalizăm ideile de mai sus în modul următor. Există o mulțime S de *stări ale naturii* sau *evenimente incerte*. Există de asemenea o mulțime E de *experimente* disponibile, care pot fi făcute în vederea obținerii unor informații suplimentare privind evenimentele incerte. Vom nota prin e_0 *experimentul nul*, înțelegînd prin aceasta alternativa de a nu face nici un experiment; vom presupune că e_0 este totdeauna un element al lui E. Dacă E se reduce la singurul element e_0, spunem că avem o problemă de decizie *fără informație suplimentară*.

Să presupunem acum că se efectuează un anumit experiment $e \in E$. Vom nota prin $R(e)$ mulțimea rezultatelor posibile ale acestui experiment, care mai sînt numite uneori și *mesajele* acestui experiment; evident, $R(e)$ trebuie să fie o listă exhaustivă și exclusivă a acestor rezultate. Mulțimea $R(e_0)$ a rezultatelor asociate experimentului nul se reduce la un singur element, pe care îl vom numi *rezultatul nul* sau *mesajul nul* și care va fi notat prin r_0.

După ce s-a ales un experiment $e \in E$ și s-a obținut rezultatul său $r \in R(e)$, urmează să alegem o decizie d din mulțimea D a deciziilor. În general, nu orice decizie este disponibilă în această situație. Din acest motiv, vom defini *mulțimea deciziilor disponibile* în urma efectuării experimentului $e \in E$ și a obținerii rezultatului său $r \in R(e)$ și o vom nota prin $D(e, r)$. Este clar că

$$D = \bigcup_{e \in E} \bigcup_{r \in R(e)} D(e, r).$$

Dacă se adoptă $d \in D(e, r)$ și starea adevărată a naturii este $s \in S$, rezultă o anumită consecință $c(e, r, d, s)$; mulțimea acestor consecințe va fi notată prin C. Este util în multe cazuri să gîndim decizia d nu ca un element al mulțimii $D(e, r)$ ci ca o funcție definită pe S și cu valori în C.

Unei probleme de decizie specificată de elementele introduse mai sus i se poate atașa un *arbore de decizie*, care descrie, în ordine cronologică, modul de acțiune al organului de decizie

și al „naturii": decidentul alege $e \in E$, „natura" produce rezultatul $r \in R(e)$, decidentul alege $d \in D(e, r)$, natura produce „starea" $s \in S$ și apoi apare consecința $c(e, r, d, s)$ din mulțimea C. Nodurile arborelui sînt asociate cu momentele de timp în care decidentul sau natura face o alegere, iar muchiile care pleacă dintr-un nod sînt asociate alternativelor deschise decidentului sau naturii în acel moment.

Pentru a ilustra modul de construire al arborelui de decizie, să considerăm următorul exemplu simplu. Avem:

$$S = \{s_1, s_2, s_3\}; \quad E = \{e_0, e_1, e_2\},$$

$$R(e_0) = \{r_0\}; \; R(e_1) = \{r_1, r_2\}; \; R(e_2) = \{r_1, r_3\};$$

$$D(e_0, r_0) = \{d_1, d_2\}; \; D(e_1, r_1) = \{d_1, d_3\};$$

$$D(e_1, r_2) = \{d_2, d_4\}; \; D(e_2, r_1) = \{d_1, d_4\}; \; D(e_2, r_3) = \{d_3, d_4\};$$

$$C = \{c_1, c_2, c_3, c_4, c_5\},$$

iar funcțiile $d_i : S \to C$ sînt date în tabelul 1.3.1. Arborele de decizie corespunzător este dat în fig. 1.3.1.

Tabelul 1.3.1

S \ D	d_1	d_2	d_3	d_4
s_1	c_1	c_2	c_3	c_4
s_2	c_3	c_4	c_1	c_2
s_3	c_2	c_5	c_2	c_3

În acest arbore de decizie, nodurile corespunzătoare decidentului au fost însemnate cu un pătrat, iar cele corespunzătoare „naturii" cu un cerc, convenție ce va fi păstrată în întreaga lucrare.

Muchiilor care pleacă din nodurile corespunzătoare naturii li se vor asocia probabilitățile de realizare ale acestor alternative, condiționate de informația disponibilă la momentul în care acestea pot să apară. Preferințele privind consecințele se vor reflecta într-o *funcție de utilitate* $u : C \to R$, unde R este mulțimea numerelor reale. Alegerea experimentelor și

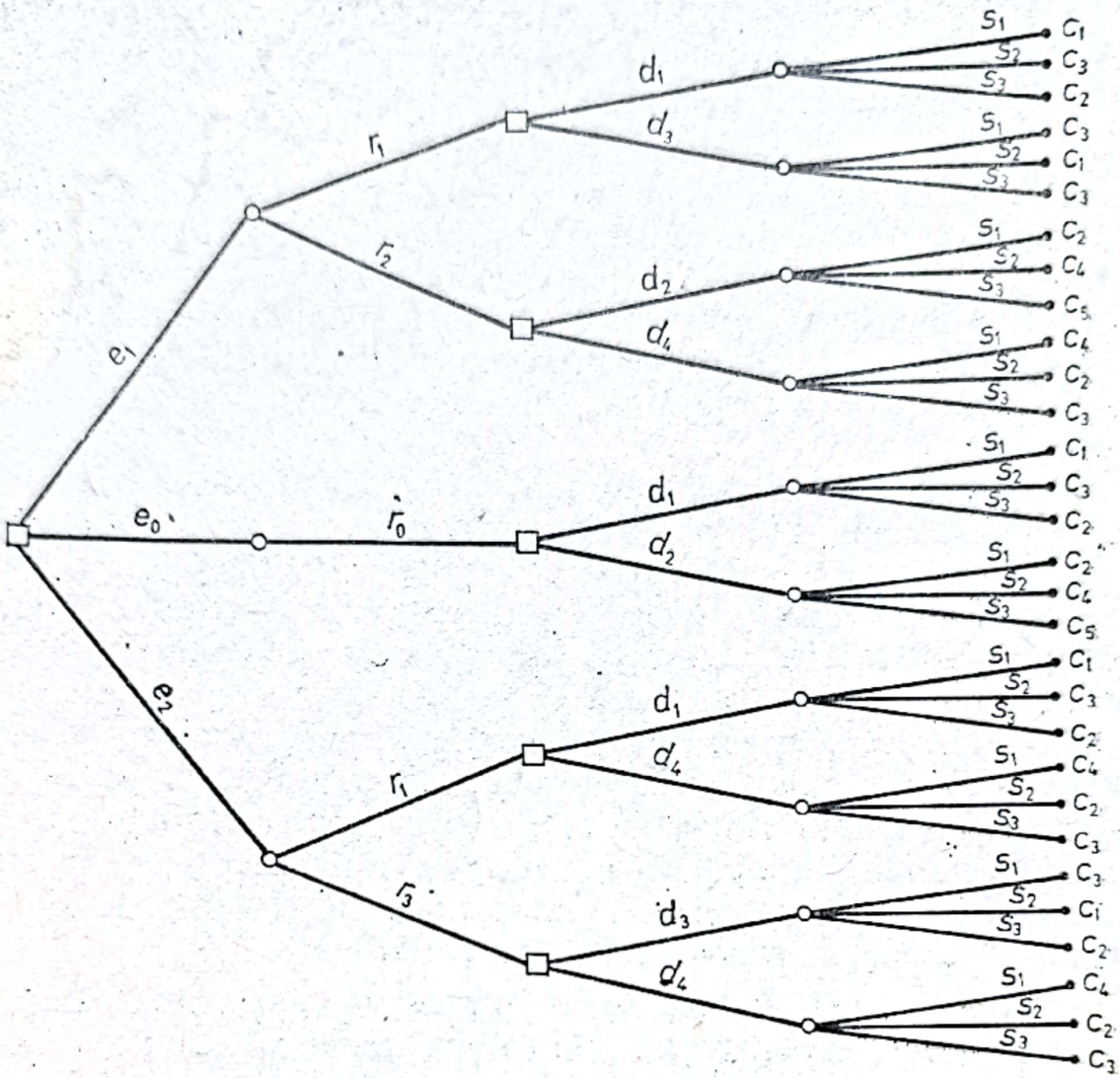


Fig. 1.3.1

deciziilor se va face astfel încît *valoarea medie a utilității* asociate să fie maximă.

Rezolvarea problemei se face în două etape. Pentru fiecare experiment $e \in E$ și rezultat obținut $r \in R(e)$ se determină mai întîi o *decizie optimă* $d^*(e, r) \in D(e, r)$ și apoi o alternativă optimă de experimentare $e^* \in E$. Implementarea acestei soluții se face în ordine inversă: se face mai întîi experimentul $e^* \in E$, se observă rezultatul său $r^* \in R(e^*)$ și apoi se ia decizia optimă $d^* = d^*(e^*, r^*)$.

Problema de decizie formulată mai sus este „într-o singură etapă". Putem considera probleme de decizie în mai multe etape, fiecare etapă avînd structura prezentată mai sus; vom reveni asupra acestei probleme în ultima parte a lucrării.

Scanned with OKEN Scanner

1.4. OBSERVAȚII GENERALE PRIVIND REZOLVAREA UNEI PROBLEME DE DECIZIE

Așa cum am mai remarcat, procesul de luare a deciziilor se află în centrul celor mai multe probleme din diverse domenii ale activității umane. Analiza deciziilor propusă mai sus este un procedeu logic de evaluare a factorilor care influențează o decizie. Această analiză ar putea fi structurată în trei faze principale: faza deterministă, faza probabilistă și faza experimentării.

În prima fază se definește mai întîi în mod precis ce decizie urmează să fie luată și cine este organul de decizie. Se identifică apoi alternativele, conform cu cele precizate în § 1.1. Ținînd seama de lista alternativelor, se identifică în continuare consecințele posibile ale adoptării lor, precum și factorii care le influențează. Pe această bază se identifică în continuare variabilele de stare, care sînt asociate diferiților factori incerți, și se atribuie valori consecințelor. Urmează apoi stabilirea legăturilor existente între variabilele de stare și studierea acestora în vederea stabilirii variabilelor de stare esențiale. În sfîrșit, se determină preferințele în raport cu timpul, de obicei prin indicarea unor coeficienți de actualizare.

În această primă fază se analizează de asemenea alternativele disponibile pentru a stabili dacă unele dintre ele nu pot fi eliminate din considerații, fiind evident mai proaste (vezi § 2.2 pentru precizări suplimentare în această direcție).

În cea de a doua fază se trece mai întîi la codificarea incertitudinii asupra variabilelor de stare esențiale. Aceasta se face pe baza rezultatelor teoriei probabilităților, căreia îi este consacrată partea a doua a acestei lucrări. Tot în această fază se codifică preferințele asupra riscurilor implicate de incertitudine și de preferințele privind consecințele diverselor decizii. Aceasta se face pe baza rezultatelor *teoriei utilității*, căreia îi este dedicată cea de a treia parte a acestei lucrări. Această fază se termină cu selectarea celor mai bune alternative.

În sfîrșit, în ultima fază analizăm mai întîi implicațiile pe care le-ar avea cunoașterea certă a variabilelor de stare esențiale, sau, cum se mai spune, *eliminarea incertitudinii* asupra acestor variabile de stare. *Informația perfectă* sau *clarviziunea*

Scanned with OKEN Scanner

are o *valoare*, care rezultă din diferența dintre valoarea medie a utilității obținută înainte și după eliminarea incertitudinii. Cunoscînd valoarea informației perfecte, putem aprecia dacă este necesar și cum anume trebuie îmbogățită informația disponibilă. Aceasta ne permite deci să evaluăm diverse programe de procurare a informației suplimentare (experimente) în scopul depistării unui program cît mai economic. Dacă un astfel de experiment pare să fie atrăgător, ne întoarcem la faza probabilistă, pe care o reluăm în aceste noi condiții.

Subliniem în încheiere că trebuie făcută o distincție netă între „decizii bune" și „consecințe bune". În limbajul obișnuit, o decizie este apreciată ca bună dacă, în urma adoptării sale, se obține o consecință bună. Dificultatea constă în faptul că nimeni nu știe să aleagă o astfel de decizie într-o situație incertă, deoarece consecințele apar din jocul incertitudinii și al deciziei.

Nu aceasta este accepțiunea pe care o dăm deciziilor „bune" în teoria deciziei. În cadrul acestei teorii, o decizie este considerată „bună" dacă este luată pe baza unor raționamente corecte, dacă este deci coerentă cu aprecierile asupra incertitudinii și preferințele asupra consecințelor. Evident, există posibilitatea ca decizii bune în acest sens să aibă consecințe proaste. E adevărat, șansele de apariție ale consecințelor proaste sînt în general mici, dar posibilitatea unei consecințe proaste nu este exclusă.

Aceasta ar putea fi considerată o parte slabă a acestei teorii. Nu cunoaștem însă alte moduri de rezolvare ale problemei de decizie care să fie superioare în această privință teoriei deciziei. Dimpotrivă, avem toate motivele să credem că orice altă abordare este inferioară din acest punct de vedere teoriei deciziei.

1.5. SCURT ISTORIC

După cum am văzut mai sus, teoria probabilităților și teoria utilității sînt ingredientele de bază care intervin în formularea și rezolvarea problemelor de decizie. Vom prezenta în cele ce urmează o scurtă schiță a dezvoltării și a situației actuale ale acestor teorii.

Scanned with OKEN Scanner

Noțiunea de probabilitate pare să fi fost introdusă prima oară de Gerolamo Cardano (1501—1576) în legătură cu jocurile de noroc, deși interesul pentru hazard pare să fi fost tot atît de vechi ca și celelalte preocupări umane. Cardano a definit probabilitatea ca raportul dintre numărul cazurilor favorabile și numărul tuturor cazurilor egal probabile, în legătură cu preocupările sale privind aruncarea zarurilor. Surprinderea legăturii dintre probabilitatea astfel definită și frecvența relativă de producere a unei anumite situații într-un mare număr de experimente „identice" pare să fi fost realizată pe deplin de Denis Poisson în 1837.

În ambele situații, probabilitatea este privită ca un parametru fizic, cum ar fi, de exemplu, masa sau volumul, fiind deci asociată cu starea reală a lucrurilor; în primul caz este un raport de cazuri descrise fizic, iar în al doilea este limita frecvenței relative, care este tot un concept fizic.

În 1713, în lucrarea sa *Ars Conjectandi*, Jacques Bernoulli aduce un punct de vedere diferit în legătură cu conceptul de probabilitate. J. Bernoulli sugerează că probabilitatea este un „grad de încredere" sau „un grad de convingere" pe care un individ îl asociază cu un anumit eveniment incert și care depinde de informațiile sale asupra acestui eveniment; evident, acest grad de convingere poate să difere de la persoană la persoană, ceea ce explică denumirea de *probabilitate subiectivă* pentru această noțiune. Pe aceeași linie subiectivă, dar cu o nuanță diferită, se pot menționa Pierre Simon de Laplace și Augustus de Morgan. Primul afirma în lucrarea sa *Essai philosophique sur les Probabilités* că probabilitatea nu este decît „expresia ignoranței omului". Al doilea afirma în lucrarea sa *Formal Logique* (1847) că „prin grad de probabilitate vrem să spunem în realitate, sau ar trebui să spunem, grad de încredere".

Aceste idei sînt continuate și îmbogățite cu nuanțe noi la începutul secolului nostru. Astfel, R. L. Keynes, în lucrarea sa *Treatise on Probability* (1921), susține teza conform căreia probabilitatea exprimă un „grad de încredere rațional (logic)", care leagă în mod logic un ansamblu de propoziții, luate ca ipoteze, cu o altă propoziție, care reprezintă concluzia. Acestui punct de vedere i se asociază și Jeffreys în lucrarea sa *Theory of Probability* (1939).

În lucrarea sa, *Truth and Probability* (1926), Ramsey adoptă punctul de vedere subiectiv așa cum este cunoscut și în prezent.

Scanned with OKEN Scanner

Mai precis, în această viziune se apreciază că probabilitățile nu trebuie considerate ca expresii ale unui grad de încredere logic, ci ca expresii ale unui grad de încredere subiectiv, interpretat ca fiind semnificativ pentru comportamentul individului pus în fața unor pariuri simple, evidente. Pe aceeași linie se înscriu contribuțiile lui Bruno de Finetti și E. Borel, consacrate aceleiași problematici. Plecînd de la nevoia de coerență în aceste comportamente, de Finetti demonstrează că probabilitățile subiective satisfac legile clasice ale teoriei probabilităților.

Teoria utilității, așa cum este cunoscută astăzi, a fost dezvoltată într-un caz particular de von Neumann și Morgenstern în lucrarea lor *Theory of games and Economic behaviour* (1947). Contribuții anterioare au mai fost aduse în acest domeniu de D. Bernoulli, G. Cramer și F. Ramsey.

O contribuție considerabilă la construirea unei teorii a deciziei în condiții de incertitudine a adus-o Abraham Wald, contribuție expusă magistral în lucrarea sa *Statistical Decision Functions* (1950). Continuînd tradiția școlii de statistică matematică fondată de Neyman și Pearson, Wald formulează problemele fundamentale ale statisticii matematice ca *probleme de decizie*. Generalizînd problemele de estimație punctuală, estimații prin regiuni de încredere și testarea ipotezelor statistice, Wald ajunge la formularea unei probleme de decizie care este, în esență, cea descrisă în § 1.3. Pentru fiecare stare a naturii și pentru fiecare alternativă de experimentare și acțiune se determină valoarea (utilitatea) asociată. Se delimitează apoi mulțimea alternativelor *eficiente* (sau *admisibile*, cum le numește Wald), mulțime care conține toate alternativele interesante în problema dată (vezi § 2.2 pentru detalii). Din păcate, aceste mulțimi eficiente s-au dovedit prea bogate pentru clasele uzuale de probleme de decizie și o alegere finală era încă necesară. Confruntat el însuși cu această problemă, Wald era de părere că alegerea finală este o chestiune de apreciere, exterioară problemei de decizie, și nu poate fi tranșată de o manieră științifică.

Practicianul are însă nevoie de criterii clare și de o soluție unică la problema sa. Pornindu-se de la această nevoie de unicitate, s-au încercat diverse modalități de a selecta „în mod obiectiv" o alternativă din mulțimea eficientă obținută. Diversele puncte de vedere în această direcție au fost rînd pe rînd puse în dificultate de contraexemple care arătau inadec-

Scanned with OKEN Scanner

varea unuia sau a altuia dintre criteriile propuse pentru selecția finală.

O serie de cercetători în acest domeniu au avut ideea să procedeze invers, adică să precizeze caracteristicile pe care ar trebui să le aibă orice principiu „rezonabil". Aceasta a condus la introducerea teoriei utilității și a probabilităților subiective asociate stărilor naturii, dînd teoriei deciziei în condiții de incertitudine forma sub care este cunoscută în prezent. Contribuții majore în această direcție au adus L. J. Savage, H. Chernoff, K. Arrow, H. Raiffa, D. Luce etc. Monografiile lui Savage (1954 și 1962) și Luce și Raiffa (1957) conțin principalele idei în această direcție.

În ultimele două decenii, eforturile teoretice s-au concentrat pe de o parte pe probleme de codificare a informației, iar pe de altă parte pe elaborarea unei teorii a utilității care să poată fi utilizată cu succes în atacarea problemelor de decizie complexe cu care este confruntată societatea modernă. O serie de aplicații ale acestor idei la atacarea unor probleme de decizie s-au dovedit utile și promițătoare pentru viitoarele încercări. Există puncte ale teoriei încă incomplet elaborate, iar „tehnologia" de aplicare a acestor idei generale este încă destul de săracă pentru problemele reale cu care sîntem confruntați. Eforturile intense care se depun actualmente în această direcție, precum și interesul crescînd pentru aceste metode reprezintă garanția unor succese mari în viitorul apropiat.

1.6. PRO ȘI CONTRA ACESTEI TEORII

Încă de la început trebuie să atragem atenția că teoria deciziei în condiții de incertitudine este o teorie *normativă*, indicînd modul în care ar trebui acționat pentru a realiza o comportare rațională, coerentă cu preferințele și aprecierile decidentului privind incertitudinile și consecințele actelor sale. Există numeroase dovezi care atestă că multe organe de decizie nu se comportă coerent. Aceasta înseamnă că teoria deciziei are o valoare redusă în explicarea și predicția comportării unui decident tradițional.

Scanned with OKEN Scanner

Pornind de la această împrejurare, o serie de critici neagă importanța acestei teorii pentru acțiunea umană practică, opunîndu-i ceea ce s-a numit *modelul cibernetic de decizie**. Se observă mai întîi, în mod îndreptățit, că, în abordarea normativă, se cere ca organul de decizie să recurgă la calcule deosebit de complexe pentru depistarea unui mod adecvat de acțiune. Deoarece multe organe de decizie reușesc să ajungă la o decizie adecvată fără asemenea calcule se trage concluzia că cel puțin o parte dintre procesele de decizie trebuie să se desfășoare altfel.

Prima ipoteză care stă la baza modelului cibernetic de decizie constă tocmai în presupunerea că organul de decizie tinde să elimine calculul direct al consecințelor și deci să elimine impactul incertitudinii. O a doua ipoteză constă în esență în presupunerea că esența procesului de decizie este efectuarea unui set de procedee și nu alegerea unei acțiuni în acord cu preferințele. În sfîrșit, ultima ipoteză este că organul de decizie este senzitiv la o gamă foarte limitată de informații, definită ca fiind informația pertinentă în raport cu setul de proceduri stabilite.

Procedurile efectuate de un organ de decizie „cibernetic" produc consecințe, dar acestea nu intră în intenția sa. În cazul în care organul de decizie operează cu succes, „mediul înconjurător" este structurat astfel încît să mijlocească o conexiune inversă între efectele produse de procesul de decizie și informația necesară în etapele următoare.

Este clar că un proces de decizie care operează în acest mod nu presupune existența unui criteriu precis de acțiune și nu cere raționamente probabiliste, adică neagă principalele ingrediente ale modelului normativ al procesului de decizie.

În legătură cu meritele celor două abordări ale problemei de decizie, trebuie să recunoaștem relevanța modelului cibernetic pentru explicarea și predicția comportamentului unei anumite categorii de organe de decizie. Paradigma cibernetică poate de asemenea să se dovedească mai adecvată decît cea normativă în unele probleme de decizie în care mecanismele

---

* Această abordare mai este numită și *paradigma cibernetică*, opusă *paradigmei analitice*, care este prezentată în această lucrare. Pentru detalii vezi J. STEINBRUNER (1974).

Scanned with OKEN Scanner

de decizie sînt incluse în sisteme mai cuprinzătoare, cu structură cunoscută.

Sîntem însă absolut convinși că există o gamă foarte largă de procese de decizie în care teoria normativă introdusă mai sus este singura relevantă. Faptul că ea este normativă și nu descriptivă trebuie considerat mai curînd ca un avantaj. Într-adevăr, dacă organele de decizie s-ar comporta coerent, n-ar mai fi nevoie de o teorie și de mijloace științifice pentru a ajunge la coerență.

Un alt avantaj al teoriei deciziei este faptul că metodologia analizei obligă organul de decizie să privească problema ca un tot organic, să surprindă legăturile dintre diferitele aspecte ale problemei sale. Aceasta favorizează comunicarea între specialiștii de diverse profile (ingineri, economiști, medici, matematicieni etc.) care trebuie antrenați în rezolvarea diverselor părți ale problemei; formarea unui limbaj comun al acestor specialiști este unul dintre avantajele însemnate ale analizei deciziilor.

Fiecare dintre fazele de formulare și rezolvare a problemei de decizie, prezentate pe larg în § 1.4, își poate aduce contribuția la mai buna înțelegere a acesteia: evaluarea clarviziunii poate sugera importanța obținerii de informații noi, stabilirea alternativelor poate pune în evidență idei noi, deosebit de valoroase etc.

Analiza deciziilor poate furniza de asemenea argumente pentru susținerea alternativei adoptate, pentru justificarea măsurilor luate.

Un alt avantaj al acestei teorii este că oferă un limbaj comun și un teren pe care se pot obține compromisuri între punctele de vedere diferite ale experților asupra unuia sau altuia dintre aspectele problemei; asupra acestei probleme vom reveni în § 1.7.

Există desigur și aspecte mai puțin invulnerabile ale acestei teorii. Metodologia analizei deciziilor cere descompunerea problemei complexe în probleme simple și sintetizarea logică a rezultatelor obținute în vederea construirii unui program de acțiune pentru problema complexă. Analiza problemelor simple care rezultă în acest mod cere din partea decidentului și a consilierilor săi răspunsuri clare, neechivoce, la probleme elementare, ipotetice. Dificultatea, care poate părea paradoxală la prima vedere, este că aceste probleme elementare sînt foarte greu de

Scanned with OKEN Scanner

rezolvat, necesitînd explicitarea punctelor de vedere pînă în cele mai mici detalii. Din acest motiv, mulți decidenți preferă deseori refugiul unei probleme complexe, coerența comportamentelor fiind foarte greu de urmărit în aceste cazuri.

Una dintre criticile frecvent aduse acestei teorii este următoarea. Ambiția analizei deciziilor este să cuantifice incertitudinea și preferințele pentru consecințe în toate situațiile. Obiecția este că, pentru multe situații complexe reale, această ambiție este nerealistă. Mai mult, aspectele care nu pot fi cuantificate ușor sînt lăsate în afara analizei. Există desigur motive întemeiate pentru astfel de suspiciuni; unele lipsuri în elaborarea teoriei utilității, ca și lipsa de experiență suficientă în aplicarea acesteia pentru cuantificarea preferințelor sînt în mare măsură responsabile în această direcție. Totuși, trebuie să menționăm că eforturile din ultimii ani au contribuit substanțial la rezolvarea în mare măsură a acestor probleme *.

În sfîrșit, numeroase critici aduse teoriei se referă la faptul că nevoia de a cuantifica, combinată cu multitudinea factorilor care trebuie luați în considerare, conduce la abandonarea unor factori calitativi, mai greu cuantificabili, dar care pot juca rolul esențial în aprecierea unei situații. Trebuie să recunoaștem că acest lucru se poate întîmpla și critica este în unele cazuri justificată. Totul depinde de abilitatea analistului, de colaborarea sa cu organul de decizie și cu experții săi, dar, mai ales, depinde de dezvoltarea unor tehnici corespunzătoare de măsurare și a unui sistem de indicatori relevanți pentru situația dată.

O ultimă problemă care trebuie lămurită este dacă, pentru o anumită problemă de decizie, merită să facem o analiză formală sau nu. Cu alte cuvinte, avem de analizat „decizia de a face analiza deciziilor".

Gama problemelor de decizie este extrem de largă, de la decizii mărunte, cotidiene, pe care le luăm fiecare în parte, pînă la decizii de importanță majoră, cu care sînt confruntate organele de decizie de nivel foarte înalt.

Răspunsul la problema pusă mai sus este greu de dat în general, deși unele situații concrete se pot trata fără dificul-

---

* Pentru o discuție detaliată vezi R. L. KEENEY, H. RAIFFA (1976).

tate. Dacă, de exemplu, problema de decizie implică alocarea unei sume de ordinul milioanelor sau zecilor de milioane de lei, decizia de a cheltui cîteva mii de lei pentru o analiză sistematică a problemei este de bun simț. Unii teoreticieni avansează chiar ideea că 1% din valoarea asociată deciziei ar trebui cheltuită pentru analiza sistematică a problemei.

1.7. NECESITATEA UNEI METATEORII

Una dintre cele mai simple probleme de decizie este adoptarea sau respingerea unei ipoteze, care poate fi în realitate adevărată sau falsă. Astfel de probleme apar, de exemplu, în acceptarea sau respingerea loturilor de produse de calitate necunoscută (vezi și § 1.2.1, (e)). Dacă se respinge ipoteza cînd ea este adevărată se comite o *eroare de primul tip*, iar dacă se acceptă ipoteza cînd ea este falsă se comite o *eroare de al doilea tip*.

Putem să gîndim și la o *eroare de al treilea tip:* rezolvarea unei false probleme*. Se apreciază că există mulți practicieni care, cel puțin la începuturile activității lor, rezolvă false probleme.

Una dintre sursele acestei erori este slaba legătură dintre analist și organul de decizie. Pentru ca analistul să rezolve adevărata problemă este necesar să existe o comunicare stabilă, în ambele sensuri, între el și decident, în scopul verificării permanente a faptului că s-a delimitat adevărata problemă care trebuie rezolvată.

Problema este în esență mai complicată, deoarece nu există reguli clare care să ne indice că problema pe care am ales s-o rezolvăm este cea adevărată. De exemplu, în problema cea mai simplă pe care ne-o putem imagina și care a fost menționată mai sus, alegerea unui anumit risc de primul tip α și a unui anumit risc de al doilea tip β este exterioară problemei de decizie. Din punct de vedere logic, α și β pot fi arbitrare și selectarea unor valori pentru aceste riscuri se face pe consi-

* Vezi H. RAIFFA (1970).

derente extralogice, exterioare teoriei. Mai mult, alegerea unei anumite ipoteze pentru a fi testată în primul rînd este făcută de asemenea în afara acestei teorii. Este deci de așteptat ca riscul de al treilea tip să pondereze într-un anumit mod riscurile de primul și al doilea tip. În concluzie, multe dintre obiectivele de bază ale teoriei deciziei nu pot fi atinse din interiorul acestei teorii și necesită considerente exterioare, adică necesită crearea unei metateorii.

Obiectivele unei metateorii ar putea fi pe scurt formulate în modul următor. Metateoria deciziei trebuie: (1) să ne ofere, la nivel global, criterii de alegere a problemei pe care o rezolvăm; (2) să specifice, la nivel de detaliu, structura problemei pe care am ales s-o rezolvăm și (3) să precizeze criteriile pe baza cărora facem selecția unei alternative.

O încercare de construire a unei astfel de teorii (vezi I. Mitroff, Betz, F. (1972)) pleacă de la premisa că diferența dintre aprecierile a doi experți privind aceeași problemă de decizie își are originea în diferențe de ordin mai adînc, ținînd de anumite ipoteze și principii filozofice de bază (numite, cu un singur cuvînt, *Weltanschauungen)* ale celor doi experți. Confruntarea părerilor, dezbaterea dialectică între cei doi experți au loc deci, în fapt, în domeniul principiilor de bază. Se poate construi pe aceste premise o metaproblemă de decizie, a cărei analiză furnizează elementele necesare pentru atingerea celor trei obiective descrise mai sus. Ideile prezentate mai sus sînt încă departe de a se constitui într-o metateorie a deciziei închegată, deși ele pot fi utile în anumite cazuri.

Scanned with OKEN Scanner

Capitolul 2

UN EXEMPLU

2.1. INTRODUCERE

Vom prezenta un exemplu simplu de problemă de decizie pentru a ilustra cîteva dintre aspectele generale discutate în capitolul precedent. Pentru a evita orice comentarii și ipoteze, discutabile în acest moment inițial, exemplul este de natură pur didactică*. În ultima parte a lucrării noastre, consacrată în întregime teoriei deciziei, vom avea ocazia să prezentăm și exemple de analize ale unor probleme reale.

Problema de decizie cu care ne ocupăm aici este următoarea. Avem un anumit număr de urne de două tipuri: *urne de primul tip*, care conțin fiecare 3 bile albe și 7 bile roșii, și *urne de al doilea tip*, care conțin fiecare 8 bile albe și 2 bile roșii. Se alege o urnă oarecare și ni se propune jocul următor: dacă ghicim tipul urnei, atunci cîștigăm 100 de lei dacă urna este de primul tip și 200 lei dacă urna este de al doilea tip. În cazul cînd nu ghicim tipul urnei, pierdem 30 lei dacă urna este de primul tip și 10 lei dacă urna este de al doilea tip. Ce vom face?

Conform cu planul general de acțiune, expus în §§ 1.3 și 1.4, ne vom ocupa pe rînd de diversele aspecte ale acestei probleme de decizie. Mai întîi trebuie stabilit cine este decidentul. Cei mai mulți oameni gîndesc la problemele de decizie ca spectatori pasivi, ca și cum acest lucru ar fi posibil. Sperăm ca cititorul să accepte, alături de noi, rolul decidentului; vom examina împreună, din punctul nostru de vedere, această problemă.

Un al doilea aspect care trebuie lămurit este construirea listei de decizii. Bineînțeles, putem refuza să participăm la jocul propus. Rezultă că sînt posibile următoarele decizii: (d_1) apreciem că urna este de primul tip, (d_2) apreciem că urna

---

* În analiza acestei probleme sînt necesare cîteva cunoștințe elementare de calculul probabilităților, printre care și formula lui Bayes, pentru care cititorul poate consulta § 3.5.

Scanned with OKEN Scanner

este de al doilea tip, (d_3) refuzăm să participăm la joc. Mulțimea deciziilor disponibile este deci $D = \{d_1, d_2, d_3\}$.

Consecințele la care se poate ajunge în urma adoptării uneia sau alteia dintre deciziile disponibile sînt următoarele: (c_1) cîștigarea sumei de 200 lei, (c_2) cîștigarea sumei de 100 lei, (c_3) *status quo*, (c_4) pierderea sumei de 10 lei, (c_5) pierderea sumei de 30 lei; consecința c_3 apare evident ca urmare a deciziei d_3, indiferent de starea adevărată a naturii. Mulțimea consecințelor este deci $C = \{c_1, c_2, c_3, c_4, c_5\}$. Este clar că nu există nici un dubiu în legătură cu preferințele noastre pentru aceste consecințe: c_1 este preferată lui c_2, aceasta lui c_3, aceasta lui c_4 ș.a.m.d.

În acord cu notațiile din § 1.3, vom nota prin S mulțimea stărilor naturii în acest caz. Este clar că această mulțime are două elemente, și anume: (s_1) urna este de primul tip și (s_2) urna este de al doilea tip. Rezultă deci că $S = \{s_1, s_2\}$.

Vom nota prin $c(d_i, s_j)$ consecința* care apare dacă se ia decizia d_i și starea naturii este s_j. Se constată ușor că avem:

$$c(d_1, s_1) = c_2, \quad c(d_2, s_1) = c_4, \quad c(d_3, s_1) = c_3,$$

$$c(d_1, s_2) = c_5, \quad c(d_2, s_2) = c_1, \quad c(d_3, s_2) = c_3.$$

Să ne concentrăm acum atenția asupra incertitudinii din problema noastră. Ceea ce nu știm este tipul urnei care ni se prezintă spre recunoaștere. Ceea ce știm este că ea provine dintr-o anumită mulțime de urne, dar numărul urnelor de primul și de al doilea tip în această mulțime ne este necunoscut. Deoarece urna este aleasă la întîmplare din această mulțime de urne, ar fi foarte important să știm care este proporția urnelor de primul tip în mulțimea tuturor urnelor; deocamdată presupunem că nu avem nici o informație despre această proporție și încercăm să vedem ce concluzii se pot obține în această situație de completă ignoranță.

---

* Deoarece, deocamdată, nu considerăm nici un fel de experiment, notația generală $c(e, r, d, s)$ se reduce la cea utilizată aici.

Scanned with OKEN Scanner

2.2. ANALIZA PRELIMINARĂ A PROBLEMEI

Vom nota prin $v(d_i, s_j)$ valoarea consecinței $c(d_i, s_j)$, adică valoarea asociată deciziei d_i dacă adevărata stare a naturii este s_j, definită ca fiind valoarea cîștigului asociat consecinței $c(d_i, s_j)$. Avem deci:

$$v(d_1, s_1) = 100\,;\ v(d_2, s_1) = -30\,;\ v(d_3, s_1) = 0\,;$$

$$v(d_1, s_2) = -10\,;\ v(d_2, s_2) = 200\,;\ v(d_3, s_2) = 0.$$

Spunem că o decizie d_i *domină* decizia d_j dacă

$$v(d_i, s_k) \geqslant v(d_j, s_k)$$

pentru orice k și există un indice k_0 pentru care avem

$$v(d_i, s_{k_0}) > v(d_j, s_{k_0}).$$

O decizie este numită *eficientă* sau *nedominată* dacă nu există o altă decizie care să o domine. Este clar că, într-o problemă de decizie, nu ne interesează decît deciziile eficiente, cele dominate trebuind să fie eliminate.

Se observă fără dificultate că în exemplul nostru toate cele trei decizii sînt eficiente. Alegerea uneia dintre aceste decizii nu se poate face pe baze logice, fără informații suplimentare privind incertitudinea din problemă. Există totuși cîteva puncte de vedere interesante în această direcție.

Astfel, *criteriul pesimist* al lui Wald constă în alegerea celei mai bune decizii, presupunînd că ne aflăm în cea mai defavorabilă situație. Pentru cazul nostru, cele mai mici cîștiguri asociate cu cele trei decizii disponibile sînt:

$$\min_j v(d_1, s_j) = v(d_1, s_2) = -30,$$

$$\min_j v(d_2, s_j) = v(d_2, s_1) = -10,$$

$$\min_j v(d_3, s_j) = v(d_3, s_1) = v(d_3, s_2) = 0.$$

Scanned with OKEN Scanner

Deoarece

$$\max_i \min_j v(d_i, s_j) = v(d_3, s_1) = v(d_3, s_2) = 0,$$

rezultă că cîștigul maxim garantat este egal cu zero și se obține cu decizia d_3. Decizia d_3 obținută pe această cale este *decizia pesimistă* sau *decizia max-min*, denumire justificată de ultima relație scrisă mai sus. Cu alte cuvinte, sfatul pesimistului este să nu ne angajăm în jocul propus, *dacă nu putem obține informații suplimentare* privind incertitudinea din problema noastră.

Un criteriu similar criteriului pesimist este *criteriul regretului* al lui Savage. Dacă adevărata stare a naturii este s_j, atunci

$$\rho(d_i, s_j) = \max_i v(d_i, s_j) - v(d_i, s_j)$$

are semnificația *regretului* pe care îl are decidentul care adoptă decizia d_i în locul deciziei celei mai bune în situația dată. Aplicarea criteriului pesimist valorilor $\rho(d_i, s_j)$ asociate consecințelor $c(d_i, s_j)$ reprezintă criteriul regretului al lui Savage. În exemplul nostru avem

$$\rho(d_1, s_1) = 0, \quad \rho(d_2, s_1) = 110, \quad \rho(d_3, s_1) = 100$$

și

$$\rho(d_1, s_2) = 230, \quad \rho(d_2, s_2) = 0, \quad \rho(d_3, s_2) = 200.$$

Regretele cele mai mari (pesimiste) asociate celor trei decizii sînt deci

$$\max_j \rho(d_1, s_j) = 230, \quad \max_j \rho(d_2, s_j) = 110,$$

$$\max_j \rho(d_3, s_j) = 200.$$

Deoarece

$$\min_i \max_j \rho(d_i, s_j) = \rho(d_2, s_1) = 110,$$

rezultă că cel mai mic dintre aceste regrete pesimiste este 110 lei și corespunde deciziei d_2. În cazul nostru deci, sfatul pe care ni-l dă criteriul regretului este să spunem că urna este de al

doilea tip, în speranța obținerii premiului celui mai mare, care este de 200 lei. Evident, s-ar putea ca proporția urnelor de al doilea tip să fie foarte mică și deci șansa de a obține acest premiu este foarte mică, fiind foarte probabil să pierdem 10 lei dacă urmăm acest sfat.

2.3. CONSIDERAREA UNEI REPARTIȚII APRIORI

Proporția urnelor de primul tip, pe care o vom nota prin p, și proporția q a urnelor de al doilea tip conțin evident informații de cea mai mare valoare pentru decident. Deoarece urna este selectată la întîmplare din mulțimea tuturor urnelor, rezultă că probabilitatea evenimentului incert s_1, pe care o notăm prin $P(s_1,)$ este egală cu p, iar probabilitatea evenimentului incert s_2 este egală cu q. Avem deci:

$$P(s_1) = p,\ P(s_2) = q,$$

unde

$$p + q = 1,\ p,\ q \geqslant 0.$$

O probabilitate considerată inițial pe mulțimea stărilor S este numită și *repartiție apriori*. În cazul nostru, repartiția apriori este complet determinată de valoarea lui $p \in [0,1]$.

Este aici locul să menționăm că, în multe probleme reale, o repartiție apriori este dificil de atribuit, iar majoritatea teoreticienilor (numiți și *obiectiviști* sau *clasici*) contestă chiar existența unei astfel de repartiții. Discuția din paragraful precedent asupra exemplului nostru arată cît de puțin profitabilă este analiza deciziilor fără considerarea unei repartiții apriori.

Dacă se consideră o repartiție apriori determinată de o valoare $p \in [0, 1]$ precizată, atunci putem asocia pentru problema noastră, în această etapă de studiu, arborele de decizie reprezentat în fig. 2.1.1, care nu mai necesită explicații suplimentare (vezi § 1.3), cu excepția poate a probabilităților p și $q = 1-p$ care apar pe muchiile asociate evenimentelor incerte s_1 și s_2.

Pentru simplitatea analizei, să considerăm mai întîi cazul în care $p \in [0, 1]$ este cunoscut, să spunem $p = 0,6$ (și deci

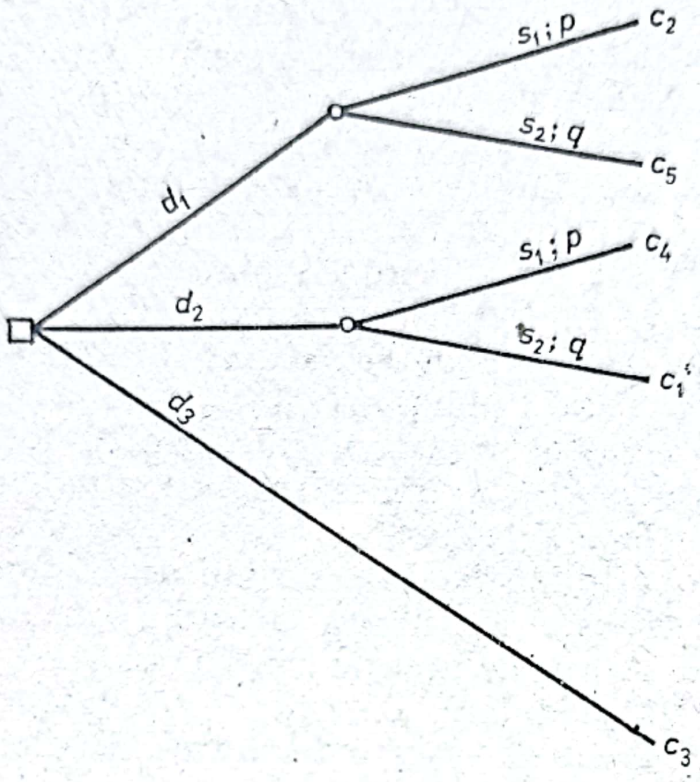


Fig. 2.1.1

$q = 0,4$). Să presupunem că alegem decizia d_1; în acest caz, cu probabilitatea 0,6 vom cîștiga 100 lei și cu probabilitatea 0,4 vom pierde 30 lei. Cu alte cuvinte, a alege d_1 este echivalent cu a alege *loteria* cu premiile 100 lei și —30 lei, obținute respectiv cu probabilitățile 0,6 și 0,4, care poate fi reprezentată schematic ca în fig. 2.1.2 (a). Analog, a opta pentru decizia d_2 este echivalent cu a opta pentru loteria prezentată în fig. 2.1.2 (b).

Compararea unor astfel de loterii este o problemă complicată, care stă la baza interesului pentru teoria utilității, asupra căreia

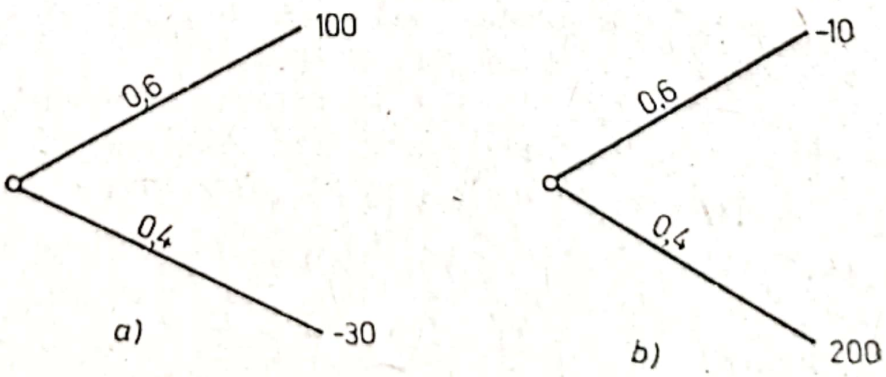


Fig. 2.1.2

vom reveni pe larg în partea a treia a acestei lucrări. Avînd în vedere caracterul ilustrativ al acestui exemplu, ne vom mărgini să introducem noțiunea de *echivalent monetar cert* (EMC) asociat unei loterii ca fiind suma în schimbul căreia deținătorul unei loterii este dispus să cedeze această loterie. Cu alte cuvinte, a decide între loterii înseamnă a determina echivalentul monetar cert al fiecăreia și a opta pentru loteria cu cel mai mare echivalent monetar cert. Modul în care un individ atribuie un echivalent monetar cert unei loterii exprimă *preferințele sale față de risc*, prudența sau, dimpotrivă, dorința de aventură. În general, un echivalent cert mai mic pentru loteriile noastre înseamnă o prudență mai mare sau, cu alte cuvinte, o mai mare *aversiune față de risc*.

Una dintre posibilitățile de a atribui un echivalent monetar cert unei loterii este asimilarea acestuia cu cîștigul mediu asociat loteriei. Atragem atenția că, în realitate, puțini decidenți fac această asimilare. Cei care procedează astfel se dovedesc a fi *indiferenți la risc;* majoritatea decidenților au însă aversiune la risc! Pentru simplitate, vom presupune că noi, decidentul, sîntem indiferenți la risc, lăsînd aspectele mai delicate pentru mai tîrziu.

Pentru loteria (a) din fig. 2.1.2, valoarea medie a cîștigului este

$$0{,}6 \cdot 100 + 0{,}4 \cdot (-30) = 4{,}80 \text{ lei},$$

iar pentru loteria (b) din aceeași figură valoarea medie a cîștigului este

$$0{,}6 \cdot (-10) + 0{,}4 \cdot 200 = 7{,}40 \text{ lei}.$$

Dacă se atribuie celor două loterii aceste echivalente monetare certe, rezultă că problema de decizie considerată mai sus este rezolvată: cea mai bună decizie este d_2, pentru care se obține un echivalent monetar cert maxim.

Revenind la cazul cînd $p \in [0, 1]$ este necunoscut, valorile medii ale cîștigurilor obținute pentru deciziile d_1, d_2, d_3 vor fi notate respectiv prin $v_1(p)$, $v_2(p)$, $v_3(p)$. Este clar că avem:

$$v_1(p) = p \cdot 100 + (1-p)(-30) = 130p - 30,$$
$$v_2(p) = p(-10) + (1-p) \cdot 200 = -210p + 200,$$
$$v_3(p) = 0.$$

Scanned with OKEN Scanner

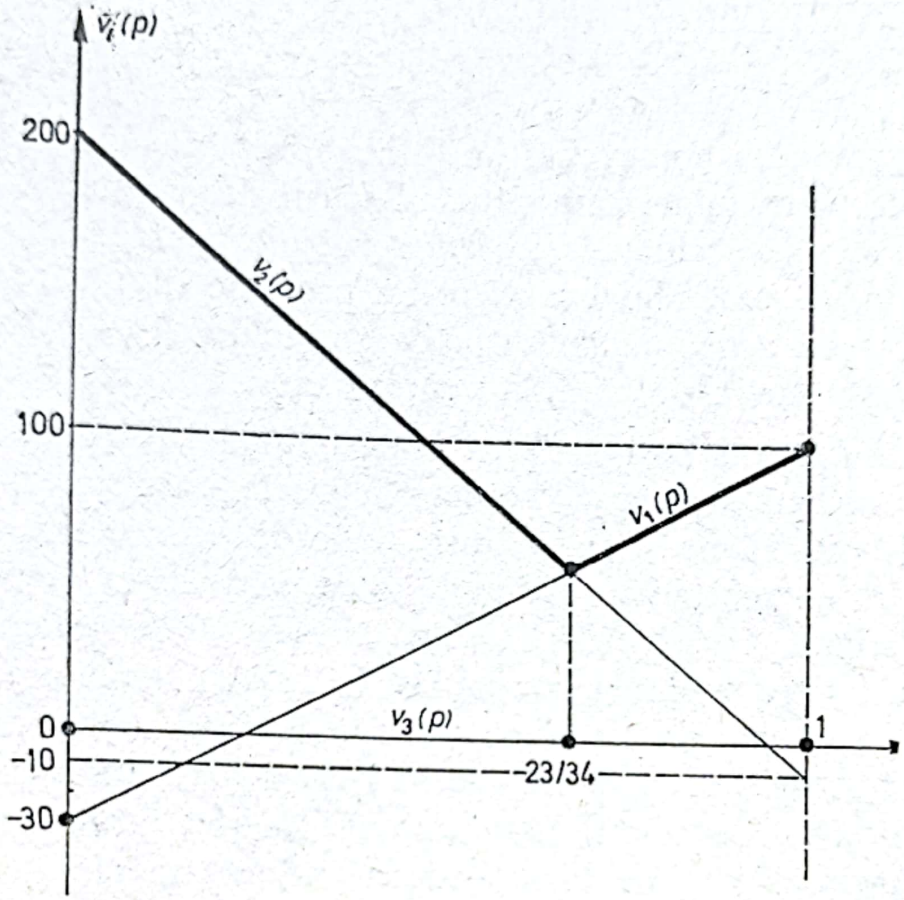


Fig. 2.1.3

În figura 2.1.3 sînt reprezentate dreptele corespunzătoare acestor funcții. Se observă ușor că decizia cea mai bună pentru $p \in [0, 23/24]$ este d_2, iar pentru $p \in [23/34, 1]$ este d_1.

Cu alte cuvinte, atît timp cît proporția urnelor de primul tip este mai mică decît $23/34 = 76\%$, vom spune că urna este de al doilea tip, iar în caz contrar vom spune că urna este de primul tip.

Pentru noi, care am optat pentru indiferență la risc, sfatul pesimist de a adopta d_3 este evident cea mai proastă alternativă, iar sfatul pe care ni-l dă criteriul regretului este valabil numai dacă proporția urnelor de primul tip nu este prea mare (nu este mai mare decît 76%). Evident, pentru alte preferințe la risc, s-ar putea ca situația să difere.

Scanned with OKEN Scanner

2.4. CLARVIZIUNE

Să considerăm acum că un binevoitor (un clarvăzător), care a aflat într-un fel sau altul adevărul despre urna asupra căreia trebuie să decidem, își oferă serviciile pentru a ne ajuta să facem o alegere. Mai exact, el ne propune divulgarea adevărului despre urnă, cu condiția ca serviciile să-i fie onorate cu o anumită sumă de bani. Acceptăm sau nu oferta sa?

Să presupunem mai întîi că $p \in [0,\ 23/34]$. În acest caz, adoptăm evident decizia d_2, conform cu cele stabilite în § 2.3. Dacă binevoitorul ne spune că urna este de primul tip, atunci ar trebui să ne schimbăm opțiunea, ceea ce ar avea drept consecință cîștigarea a 100 lei în loc de a pierde 10 lei, adică realizarea unui cîștig suplimentar de 110 lei. Dacă binevoitorul ne spune că urna este de al doilea tip, nu ne vom schimba hotărîrea luată și deci nu vom obține nici un cîștig suplimentar în această situație. Ca adepți ai indiferenței la risc, apreciem deci că informația oferită de binevoitor ne aduce un cîștig mediu suplimentar egal cu

$$p \cdot 110 + (1-p) \cdot 0 = 110\,p,$$

care este numit și *valoarea clarviziunii* sau a *informației perfecte*.

Dacă $p \in [23/34,\ 1]$, vom alege decizia d_1. Dacă aflăm că urna este de primul tip, nu ne vom schimba hotărîrea și deci nu vom cîștiga nimic în plus. Dacă însă binevoitorul ne spune că urna este de al doilea tip, atunci ne schimbăm hotărîrea, adoptînd d_2 în loc de d_1, ceea ce conduce la un cîștig suplimentar de 230 lei. Valoarea clarviziunii va fi în acest caz egală cu

$$p \cdot 0 + (1-p) \cdot 230 = -230p + 230.$$

Dacă notăm prin $\gamma(p)$ valoarea clarviziunii, avem

$$\gamma(p) = \begin{cases} 110p, & \text{dacă } p \in [0,\ 23/34], \\ -230p + 230, & \text{dacă } p \in [23/34,\ 1]. \end{cases}$$

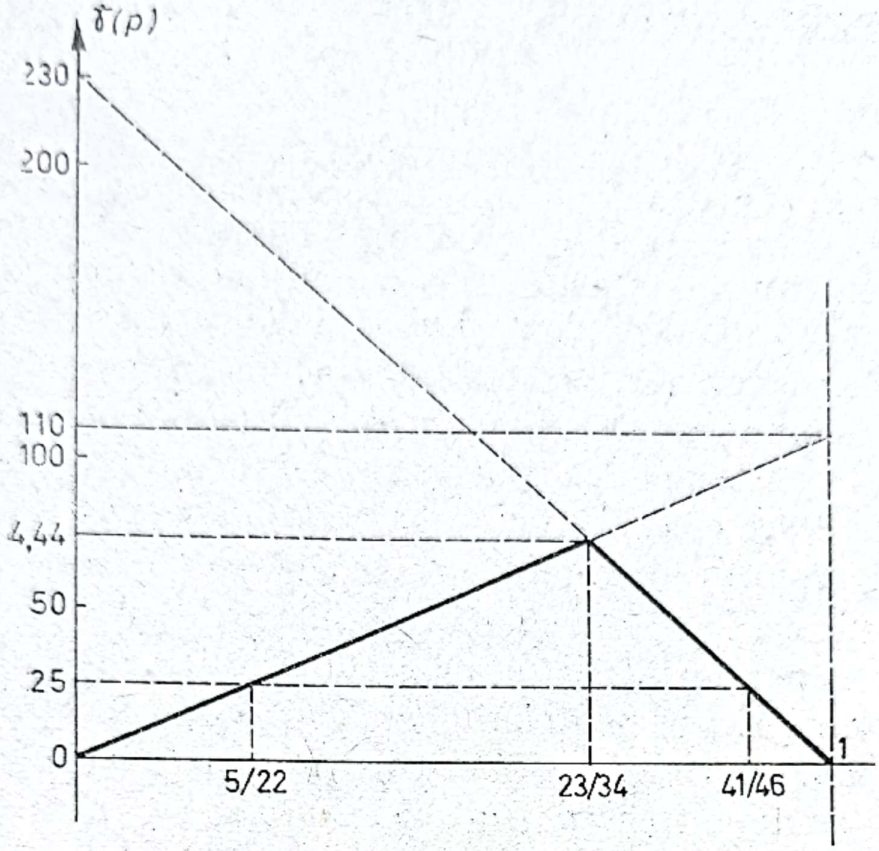


Fig. 2.1.4

Graficul valorii clarviziunii este prezentat în fig. 2.1.4.

Este clar acum că, dacă binevoitorul cere o sumă superioară sumei de 75 lei pentru informația pe care o oferă, atunci, indiferent de informația pe care o avem despre adevărata proporție a urnelor de primul tip, nu vom accepta oferta acestuia. Pe de altă parte, dacă ni se solicită 25 lei pentru aceste servicii, atunci vom accepta oferta numai în cazul în care cunoaștem că proporția urnelor de primul tip este în intervalul (5/22, 41/46).

Clarviziunea este relevantă pentru evaluarea oportunității organizării de experimente cu scopul de a obține informații suplimentare despre adevărata proporție p a urnelor de primul tip. Este clar că, dacă costul asociat experimentului depășește valoarea clarviziunii, nu are nici un sens să efectuăm aceste experimentări.

Dacă se poate însă negocia efectuarea unor experimente la un preț mai mic decît valoarea clarviziunii, s-ar putea să obținem informații valoroase, în sensul ca cîștigul mediu asociat alternativei optime să crească față de situația fără experimentare.

Scanned with OKEN Scanner

2.5. EXPERIMENTARE

Pentru a fixa ideile, să presupunem că evaluăm inițial incertitudinea în problema noastră printr-o repartiție apriori dată de valoarea $p = 0{,}5$; cu alte cuvinte, presupunem că prelevarea urnei s-a făcut dintr-o mulțime de urne conținînd un număr egal de urne de primul și al doilea tip.

Problema inițială o modificăm acum în modul următor. Avem posibilitatea să extragem din urnă o bilă contra unei sume de x lei și apoi să ghicim tipul urnei. Problema care se pune este: care este valoarea maximă a lui x pentru care experimentul propus este economic?

Notînd, ca de obicei, cu e_0 experimentul nul și cu e_1 experimentul care constă în extragerea unei bile din urnă, rezultă că mulțimea experimentelor este $E = \{e_0, e_1\}$. Notînd cu r_0 mesajul nul, iar cu r_1 și r_2 respectiv apariția unei bile albe sau roșii în urma extragerii, rezultă

$$R(e_0) = \{r_0\},\ R(e_1) = \{r_1, r_2\}.$$

Mulțimile de decizii disponibile în fiecare caz coincid cu mulțimea D. Consecințele de forma $c(e_0, r_0, d_i, s_j)$ coincid cu consecințele $c(d_i, s_j)$ din problema inițială. Consecințele asociate celorlalte situații sînt următoarele: (c_6) obținerea unui cîștig de $200-x$ lei, (c_7) obținerea unui cîștig de $100-x$ lei, (c_8) pierderea a x lei, (c_9) pierderea a $x+10$ lei, (c_{10}) pierderea a $x+30$ lei. Mai exact, avem:

$$c(e_1, r_1, d_1, s_1) = c(e_1, r_2, d_1, s_1) = c_7;$$

$$c(e_1, r_1, d_1, s_2) = c(e_1, r_2, d_1, s_2) = c_{10};$$

$$c(e_1, r_1, d_2, s_1) = c(e_1, r_2, d_2, s_1) = c_9;$$

$$c(e_1, r_1, d_2, s_2) = c(e_1, r_2, d_2, s_2) = c_6;$$

$$c(e_1, r_1, d_3, s_1) = c(e_1, r_2, d_3, s_1) = c_8;$$

$$c(e_1, r_1, d_3, s_2) = c(e_1, r_2, d_3, s_2) = c_8.$$

Scanned with OKEN Scanner

Valorile asociate consecințelor, notate prin $v(e_i, r_j, d_k, s_m)$ și care nu depind în fapt decît de e_i, d_k și s_m, sînt date în tabelul 2.1.1, care nu necesită explicații suplimentare.

Tabelul 2.1.1

	(e_0, d_1)	(e_0, d_2)	(e_0, d_3)	(e_1, d_1)	(e_1, d_2)	(e_1, d_3)
s_1	100	-10	0	$100-x$	$-(10+x)$	$-x$
s_2	-30	200	0	$-(30+x)$	$200-x$	$-x$

Se poate vedea ușor că, oricare ar fi $x > 0$, alternativa (e_1, d_3) este dominată de (e_0, d_3) și deci poate fi eliminată din considerații. Arborele de decizie corespunzător problemei pe care o considerăm în acest stadiu al analizei este prezentat în fig. 2.1.5. Pe acest arbore, în locul consecințelor am trecut direct valorile asociate lor.

Afectarea probabilităților pe acest arbore de decizie se face în modul următor. Cunoaștem probabilitățile evenimen-

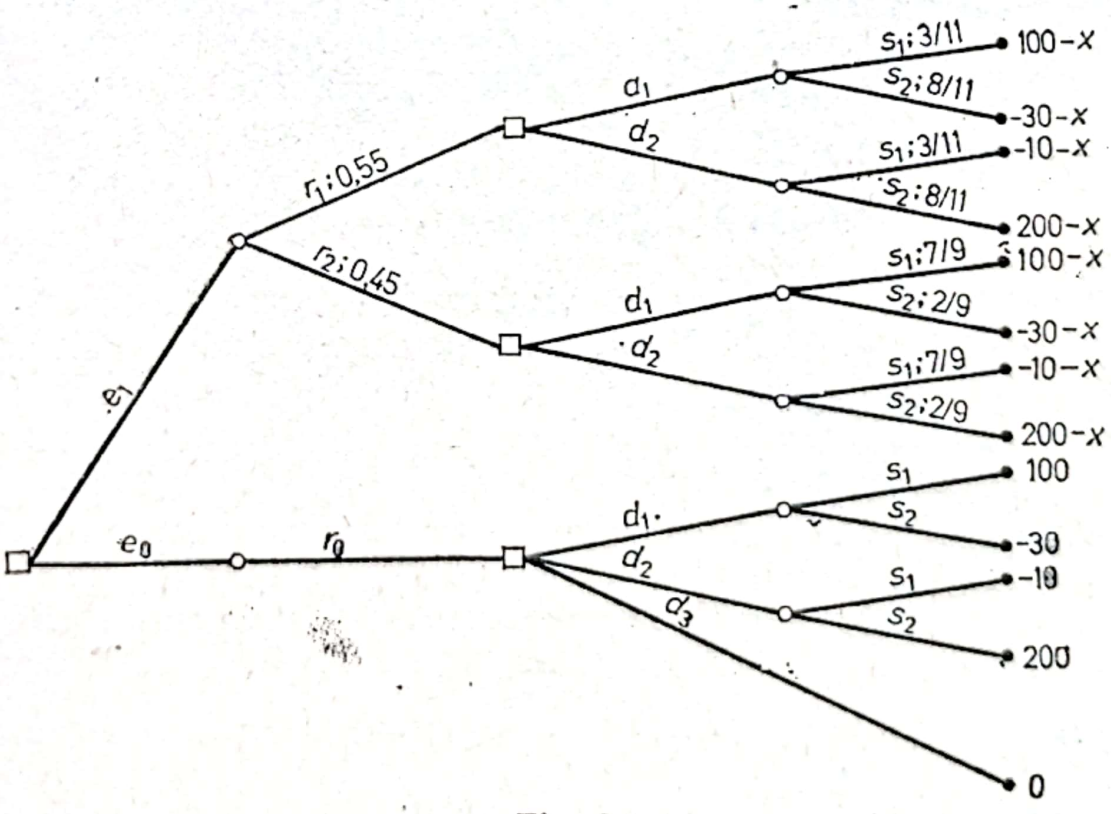


Fig. 2.1.5

telor incerte s_1 și s_2, care sînt respectiv $P(s_1) = 0,5$ și $P(s_2) = 0,5$. Dacă extragerea se face dintr-o urnă de primul tip, probabilitățile (condiționate de s_1) ale evenimentelor incerte r_1 și r_2 sînt

$$P(r_1 \mid s_1) = 3/10 = 0,3;\ P(r_2 \mid s_1) = 7/10 = 0,7.$$

Analog, dacă extragerea se face dintr-o urnă de al doilea tip, avem

$$P(r_1 \mid s_2) = 8/10 = 0,8;\ P(r_2 \mid s_2) = 2/10 = 0,2.$$

Rezultă atunci

$$P(r_1) = P(s_1)\,P(r_1 \mid s_1) + P(s_2)\,P(r_1 \mid s_2) = 0,55$$

și deci

$$P(r_2) = 1 - P(r_1) = 0,45.$$

Utilizînd teorema lui Bayes, obținem

$$P(s_1 \mid r_1) = \frac{P(s_1)\,P(r_1 \mid s_1)}{P(r_1)} = 3/11$$

și deci

$$P(s_2 \mid r_1) = 1 - P(s_1 \mid r_1) = 8/11.$$

Analog, obținem

$$P(s_1 \mid r_2) = 7/9,\ P(s_2 \mid r_2) = 2/9.$$

Probabilitățile calculate în acest mod sînt trecute pe arborele de decizie din fig. 2.1.5. Probabilitățile corespunzătoare ramurii care începe cu muchia e_0 nu au mai fost trecute, deoarece ramura respectivă este arborele de decizie considerat inițial, despre care știm deja tot ce ne interesează.

Valoarea medie a cîștigului asociat cu decizia d_i și condiționat de producerea evenimentului r_j va fi notată prin $v_i(p \mid r_j)$, $i, j = 1, 2$, unde p este proporția urnelor de primul tip; în cazul nostru $p = 0,5$. Avem deci

$$v_1(0,5 \mid r_1) = (3/11)(100-x) + (8/11)(-30-x) = -x+60/11,$$
$$v_2(0,5 \mid r_1) = (3/11)(-10-x)+(8/11)(200-x) = -x+1570/11$$

și, analog,

$$v_1(0{,}5 \mid r_2) = -x + 640/9,$$

$$v_2(0{,}5 \mid r_2) = 330/9 - x.$$

Este clar că, dacă apare o bilă albă, vom prefera d_2 lui d_1, iar dacă apare o bilă roșie, vom prefera d_1 lui d_2. Cîștigul mediu maxim asociat deciziei de experimentare, pe care îl vom nota prin $v(e_1)$, va fi deci

$$v(e_1) = 0{,}55(1570/11 - x) + 0{,}45(640/9 - x) = -x + 110{,}50.$$

Deoarece cîștigul mediu maxim pentru varianta fără experimentare, pe care îl vom nota prin $v(e_0)$, este

$$v(e_0) = v_2(0{,}5) = -210 \cdot 0{,}5 + 200 = 95,$$

rezultă că vom accepta varianta cu experimentare atît timp cît costul x al acesteia verifică inegalitatea

$$95 \leqslant 110{,}50 - x,$$

adică pentru $x \leqslant 15{,}50$ lei.

Dacă, de exemplu, în urma negocierii, obținem dreptul să extragem o bilă în schimbul sumei de 10 lei, rezultă că alternativa optimă este să procedăm la extragere și apoi să spunem că urna este de primul sau al doilea tip, după cum bila extrasă este roșie sau albă. Cîștigul mediu asociat acestui mod de acțiune este de $110{,}50 - 10 = 100{,}50$ lei și întrece cu 5,50 lei cîștigul mediu pe care l-am fi obținut fără dreptul de experimentare.

Desigur, sînt și alte experimente posibile care ar putea fi evaluate, în cazul cînd sînt negociabile. De exemplu, s-ar putea extrage două sau trei bile din urnă, sau s-ar putea extrage pe rînd cîte o bilă, cu opțiunea de a continua extragerea sau a lua una dintre deciziile d_1 sau d_2. Calculele necesare pentru aceste evaluări sînt mai complicate, dar procedeele sînt similare celor descrise mai sus.

Scanned with OKEN Scanner

Partea a doua

INCERTITUDINE

Scanned with OKEN Scanner

Capitolul 3

CÎMPURI DE PROBABILITATE

3.1. INTRODUCERE

Teoria probabilităţilor a fost iniţiată acum trei secole, în urma interesului pentru jocurile de noroc. Un scurt istoric al acestei dezvoltări a fost prezentat în § 1.5. Ca disciplină matematică, teoria probabilităţilor a fost pusă pe fundamente solide în urmă cu numai 50 ani de către A. N. Kolmogorov.

Interesul pentru această teorie matematică este dovedit şi de marele număr de matematicieni care s-au ocupat cu problemele sale. Este locul să amintim aici şi de şcoala românească de probabilităţi, iniţiată şi condusă de academicienii Octav Onicescu şi Gheorghe Mihoc, şcoală care a adus contribuţii dintre cele mai importante la dezvoltarea acestei teorii.

Deşi, ca teorie matematică, teoria probabilităţilor este acceptată în general fără rezerve, există controverse aprinse în legătură cu modul în care ea poate fi aplicată la rezolvarea problemelor practice. Mulţi utilizatori ai teoriei probabilităţilor consideră probabilitatea ca un parametru fizic, cum ar fi masa sau volumul, fiind deci asociată cu starea reală a lucrurilor. În ultimii 20 de ani, cîţiva experţi, reluînd unele puncte de vedere mai vechi, amintite succint în § 1.5, au avansat ideea că probabilitatea măsoară starea noastră de cunoaştere a lucrurilor şi nu starea lor reală. Probabilitatea este un fel de termometru mental care măsoară incertitudinea în loc de temperatură. Această idee, care nu este nouă, s-a dovedit de o importanţă esenţială în rezolvarea problemelor de decizie, aşa cum am avut ocazia să constatăm pe exemplul rezolvat în capitolul 2. Una dintre explicaţiile acestui succes este faptul că, în evaluarea probabilităţilor, se poate utiliza, în această viziune, întreaga informaţie disponibilă la un moment dat. Este aproape evident că nu este înţelept să adoptăm decizii fără a utiliza întreaga informaţie pertinentă disponibilă. Problema esenţială este însă următoarea: cum să codificăm infor-

mația disponibilă într-o formă utilizabilă? Dificultățile pe care le întîlnim în codificarea informației ca probabilități sînt numeroase. Informația disponibilă poate să varieze de la o convingere fermă, rezultată din mulți ani de experiență, pînă la o părere vagă, formată în urma cîtorva observații. Este totuși de remarcat că, în problemele care intervin în practică, nu există în general cazuri în care informația să fie complet absentă. Chiar dacă, inițial, cineva „nu are nici o idee" despre ceea ce urmează să se întîmple, totuși, dacă rezultatele care ar putea să apară sînt importante pentru individ, acesta din urmă va sfîrși prin a-și exprima starea de cunoaștere prin evaluarea probabilităților necesare.

În cele ce urmează vom prezenta elementele de teoria probabilităților necesare pentru dezvoltarea ulterioară a unei teorii a deciziei în condiții de incertitudine.

3.2. EVENIMENTE. CORPURI DE EVENIMENTE

Noțiunea de *eveniment* este una dintre noțiunile fundamentale în teoria probabilităților; noțiunile de *eveniment incert* și *eveniment cert* sau *sigur* au fost discutate pe larg în § 1.1. În cele ce urmează vom nota cu E evenimentul sigur, cu $\varnothing$ *evenimentul imposibil*, iar evenimentele incerte cu litere majuscule: A, B, ...

Să considerăm o familie de evenimente A, B, C, ... care se pot produce sau nu într-o anumită situație. Dacă producerea evenimentului A implică producerea evenimentului B, se spune că A *implică* B și se scrie $A \subset B$ sau $B \supset A$. Dacă A implică B și B implică A, spunem că evenimentele A și B sînt *echivalente* și scriem $A = B$. Este clar că orice eveniment poate fi înlocuit printr-un eveniment echivalent; în cele ce urmează nu se va face deci distincție între evenimentele echivalente.

Evenimentul care constă în producerea atît a evenimentului A cît și a evenimentului B este numit *intersecția* evenimentelor A și B și este notat prin $A \cap B$. Evenimentul care constă în producerea cel puțin a unuia dintre evenimentele A sau B este numit *reuniunea* evenimentelor A și B și este notat prin $A \cup B$.

Scanned with OKEN Scanner

Evenimentele A și B se numesc *contrare* dacă $A \cup B = E$ și $A \cap B = \varnothing$. În acest caz, B este numit *evenimentul contrar* lui A și se notează cu A^c sau CA, iar evenimentul A este contrar lui B și se notează B^c sau CB. Două evenimente se numesc *incompatibile* dacă nu pot avea loc simultan, adică $A \cap B = \varnothing$.

Definiția 3.2.1. *Spunem că o familie* $\mathfrak{D} = (A_i)_{i \in I}$ *de evenimente constituie o desfacere (partiție) a evenimentului* A *dacă sînt îndeplinite condițiile următoare:*

(D_1) I *este cel mult numărabilă;*

(D_2) *Evenimentele* A_i *și* A_j *sînt incompatibile pentru* $i \neq j$;

(D_3) $A = \bigcup_{i \in I} A_i$.

Dacă I este o mulțime finită, atunci $\mathfrak{D}$ este numită *desfacere finită*. O desfacere finită sau infinit numărabilă a evenimentului sigur E este numită și *sistem complet de evenimente*.

Exemplu. Pentru ilustrarea noțiunilor introduse mai sus să considerăm experimentul care constă în aruncarea unui zar. Vom nota prin $E_1, E_2, \ldots, E_6$ evenimentele care constau în apariția fețelor 1, 2, ..., 6 la aruncarea acestui zar; evident, toate aceste evenimente sînt întîmplătoare. Evenimentul sigur este

$$E = E_1 \cup E_2 \cup \ldots \cup E_6,$$

adică evenimentul care constă în apariția uneia dintre cele șase fețe ale zarului; este clar că evenimentele $E_1, E_2, \ldots, E_6$ formează un sistem complet de evenimente (sau o listă exhaustivă și exclusivă de evenimente incerte, cum am spus în § 1.1).

Evenimentul

$$A = E_1 \cup E_3 \cup E_5$$

este evenimentul care constă în apariția unui număr impar la aruncarea zarului, iar

$$B = E_2 \cup E_4 \cup E_6$$

este evenimentul care constă în apariția unui număr par la aruncarea zarului. Este clar că evenimentele A și B astfel definite sînt incompatibile și reprezintă, de asemenea, un sistem complet de evenimente. Se poate de asemenea vedea că evenimentul A se descompune în evenimentele E_1, E_3, E_5, iar evenimentul B în evenimentele E_2, E_4, E_6, adică $\{E_1, E_3, E_5\}$ este o desfacere a lui A, iar $\{E_2, E_4, E_6\}$ este o desfacere a lui B.

Scanned with OKEN Scanner

Evenimentele $E_1, \ldots, E_6$ nu pot fi descompuse în evenimente mai simple.

Vom spune că un eveniment este *elementar* dacă nu poate fi descompus în alte evenimente. În exemplul dat mai sus evenimentele elementare se determină fără nici o dificultate. În alte cazuri, însă, această operație nu este la fel de simplă. În construirea teoriei matematice a probabilităților se pleacă totuși de la o mulțime de evenimente elementare, numită și spațiul (mulțimea) rezultatelor unui experiment. Evenimentul sigur E este chiar acest spațiu, toate celelalte evenimente fiind submulțimi ale lui E; evenimentul imposibil $\varnothing$ este în acest caz mulțimea vidă. Dacă notăm prin $\mathcal{Q}(E)$ mulțimea părților (submulțimilor) lui E, rezultă că orice eveniment este un element din $\mathcal{Q}(E)$. Atunci cînd considerăm un experiment cu spațiul rezultatelor E, nu ne interesează, în general, toate evenimentele care se asociază experimentului ci numai o parte dintre ele; vom considera deci, în general, numai o familie $\mathcal{K} \subset \mathcal{Q}(E)$ de evenimente asociate experimentului. Pentru ca operațiile asupra evenimentelor introduse mai sus (reuniune, intersecție, negare) să nu ne conducă la evenimente din afara familiei $\mathcal{K}$, va trebui ca această familie să aibă proprietățile date în definiția următoare.

Definiția 3.2.2. *O familie* $\mathcal{K} \subset \mathcal{Q}(E)$ *se numește corp pe* E *dacă îndeplinește următoarele condiții:*

(C_1) *Pentru orice* $A \in \mathcal{K}$ *avem* $A^c \in \mathcal{K}$;

(C_2) *Pentru orice* $A \in \mathcal{K}$ *și* $B \in \mathcal{K}$ *avem* $A \cup B \in \mathcal{K}$;

(C_3) $E \in \mathcal{K}$.

Perechea $(E, \mathcal{K})$, în care $\mathcal{K}$ este un corp pe E, este numită și *cîmp de evenimente.*

Teorema 3.2.1. *Dacă* $(E, \mathcal{K})$ *este un cîmp de evenimente, atunci:*

(a) $\varnothing \in \mathcal{K}$;

(b) *Pentru orice* $A, B \in \mathcal{K}$, $A \cap B \in \mathcal{K}$;

(c) *Pentru orice* $A, B \in \mathcal{K}$, $A - B \in \mathcal{K}$;

(d) *Pentru orice* $A_1, \ldots, A_n \in \mathcal{K}$, $A_1 \cup \ldots \cup A_n \in \mathcal{K}$ *și* $A_1 \cap \ldots \cap A_n \in \mathcal{K}$.

Demonstrație. (a) Rezultă imediat din condițiile (C_1) și (C_3) ale definiției precedente. (b) Deoarece $A, B \in \mathcal{K}$, rezultă $A^c, B^c \in \mathcal{K}$ și deci

$$A \cap B = (A^c \cup B^c)^c \in \mathcal{K},$$

conform cu (C_1) și (C_2).

Scanned with OKEN Scanner

(*c*) Deoarece

$$A - B = A \cap B^c,$$

afirmația se obține din cea precedentă.

(*d*) Se face inducție după n.

Observație. Spațiul rezultatelor E asociat unui experiment poate fi constituit dintr-un număr finit sau nu de evenimente elementare. În primul caz ne putem restrînge la a aplica operațiile asupra evenimentelor de un număr finit de ori; cu alte cuvinte, în acest caz este suficient ca familia $\mathcal{K}$ să fie un corp de evenimente pe E. În al doilea caz, putem fi însă în situația de a aplica de un număr infinit de ori operațiile asupra evenimentelor, ceea ce ne poate conduce la evenimente care pot să nu mai aparțină corpului $\mathcal{K}$. Din acest motiv este necesar să introducem conceptul de *corp borelian de evenimente.*

Definiția 3.2.3. *O familie* $\mathcal{K} \subset \mathcal{P}(E)$ *este numită corp borelian pe* E *dacă îndeplinește următoarele condiții:*

(CB_1) *Pentru orice* $A \in \mathcal{K}$ *avem* $A^c \in \mathcal{K}$;

(CB_2) *Pentru orice familie cel mult numărabilă de evenimente* $(A_i)_{i \in I} \subset \mathcal{K}$ *avem* $\bigcup_{i \in I} A_i \in \mathcal{K}$;

(CB_3) $E \in \mathcal{K}$.

Perechea $(E, \mathcal{K})$, în care $\mathcal{K}$ este un corp borelian pe E, este numită *cîmp borelian de evenimente* sau *spațiu măsurabil*, ultima denumire provenind din teoria matematică a măsurii.

Se poate constata ușor că un corp borelian este corp, dar invers nu este totdeauna adevărat. În particular, rezultă că teorema 3.2.1 este valabilă și pentru un corp borelian $\mathcal{K}$. În plus se obțin proprietățile precizate în teorema următoare.

Teorema 3.2.2. *Dacă* $(A_n)_{n \in N} \subset \mathcal{K}$ *este un șir de evenimente din corpul borelian* $\mathcal{K}$ *pe* E, *atunci:*

(*a*) $\bigcap_{n \in N} A_n \in \mathcal{K}$;

(*b*) $$\limsup_n A_n = \bigcap_{n=1}^{\infty} \bigcup_{j=n}^{\infty} A_j \in \mathcal{K},$$

$$\liminf_n A_n = \bigcup_{n=1}^{\infty} \bigcap_{j=n}^{\infty} A_j \in \mathcal{K}.$$

Demonstrație. (*a*) Afirmația rezultă din relația evidentă

$$\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^c\right)^c.$$

(*b*) Afirmațiile rezultă imediat din cea precedentă.

Exemple. Să considerăm familia de evenimente $\mathcal{K}_0 = \{\varnothing, E\}$. Este ușor de văzut că $\mathcal{K}_0$ este un corp borelian pe E, numit uneori *corpul borelian banal*. De asemenea, familia tuturor evenimentelor $\mathcal{K}_t = \mathcal{P}(E)$ este tot un corp borelian pe E, numit uneori *corpul borelian total*.

Este clar că pentru orice alt corp $\mathcal{K}$ avem $\mathcal{K}_0 \subset \mathcal{K} \subset \mathcal{K}_t$.

Un alt exemplu de corp borelian rezultă din teorema următoare.

Teorema 3.2.3. *Fie $\mathcal{M} \subset \mathcal{P}(E)$ o familie de evenimente. Atunci există un corp borelian unic $\mathcal{K}$ cu proprietățile următoare:*

(*a*) $\mathcal{M} \subset \mathcal{K}$;

(*b*) *Pentru orice alt corp borelian $\mathcal{F} \supset \mathcal{M}$, avem $\mathcal{F} \supset \mathcal{K}$.*

Demonstrație. Se poate arăta fără dificultate că intersecția unei familii oarecare de corpuri boreliene pe E este tot un corp borelian pe E. Dacă notăm prin $\mathcal{S}(\mathcal{M})$ mulțimea tuturor corpurilor boreliene pe E care conțin familia $\mathcal{M}$, atunci un corp borelian $\mathcal{K}$ cu proprietățile (*a*) și (*b*) de mai sus este evident intersecția tuturor corpurilor boreliene care aparțin familiei $\mathcal{S}(\mathcal{M})$, adică:

$$\mathcal{K} = \bigcap_{\mathcal{F} \in \mathcal{S}(\mathcal{M})} \mathcal{F}.$$

Un corp borelian $\mathcal{K}$ care are proprietățile (*a*) și (*b*) este unic. Într-adevăr, dacă $\mathcal{K}_1$ și $\mathcal{K}_2$ sînt două corpuri cu aceste proprietăți, atunci, aplicînd (*a*) pentru $\mathcal{K}_1$ și (*b*) pentru $\mathcal{K}_2$, obținem $\mathcal{K}_1 \supset \mathcal{K}_2$, iar din (*a*) pentru $\mathcal{K}_2$ și (*b*) pentru $\mathcal{K}_1$ obținem $\mathcal{K}_2 \supset \mathcal{K}_1$. Rezultă $\mathcal{K}_1 = \mathcal{K}_2$ și unicitatea este demonstrată.

Corpul borelian unic $\mathcal{K}$ care are proprietățile (*a*) și (*b*) de mai sus este numit *corpul borelian generat de* $\mathcal{M}$ și va fi notat prin $\mathcal{B}(\mathcal{M})$.

Dacă se consideră desfacerea $\mathcal{D} = (A_i)_{i \in I}$ a evenimentului sigur E (pe care o vom numi, simplu, *desfacere*), atunci corpul borelian generat de $\mathcal{D}$ este de forma

$$\mathcal{B}(\mathcal{D}) = \left\{A : A = \bigcup_{i \in J} A_i,\ J \subset I\right\}.$$

Într-adevăr, fie

$$\mathcal{A} = \left\{A : A = \bigcup_{i \in J} A_i,\ J \subset I\right\}.$$

Este clar că $\mathcal{D} \subset \mathcal{B}(\mathcal{D})$ și deci $\bigcup_{i \in J} A_i \in \mathcal{B}(\mathcal{D})$, deoarece J este cel mult numărabilă, de unde rezultă $\mathcal{A} \subset \mathcal{B}(\mathcal{D})$. Invers, se vede ușor că $\mathcal{D} \subset \mathcal{A}$ și că $\mathcal{A}$ este un corp borelian pe E, de unde urmează $\mathcal{B}(\mathcal{D}) \subset \mathcal{A}$, ceea ce încheie demonstrația.

Dacă desfacerea $\mathcal{D}$ este finită, atunci corpul borelian generat $\mathcal{B}(\mathcal{D})$ este de asemenea finit, cum se vede ușor. Este adevărată și afirmația inversă: orice corp borelian finit este generat de o desfacere finită.

Se poate vedea de asemenea că, dacă $\mathcal{D} = \{A_1, \ldots, A_n\}$ este o desfacere finită, atunci corpul borelian generat $\mathcal{B}(\mathcal{D})$ este de forma

$$\mathcal{B}(\mathcal{D}) = \{\varnothing ; A_1, \ldots, A_n ; A_1 \cup A_2, \ldots, A_{n-1} \cup A_n ; \ldots ; A_1 \cup \ldots \cup A_n\}$$

și are 2^n elemente.

În particular, dacă se consideră desfacerea $\mathcal{D}_A = \{A, A^c\}$, unde $A \in \mathcal{P}(E)$, atunci corpul borelian generat de această desfacere

$$\mathcal{B}(\mathcal{D}_A) = \{\varnothing, A, A^c, E\}$$

mai este numit și *corpul borelian generat de evenimentul* A și va fi notat și prin $\mathcal{B}(A)$.

În sfîrșit, în cazul cînd E este mulțimea numerelor reale $R = (-\infty, +\infty)$, iar familia $\mathfrak{M} \subset \mathcal{P}(R)$ este de forma

$$\mathfrak{M} = \{(-\infty, a) : a \in R\},$$

corpul borelian $\mathcal{B}(\mathfrak{M})$ va fi numit *corpul borelian real* și va fi notat prin $\mathcal{B}$. Perechea $(R, \mathcal{B})$ va fi numită și *cîmpul borelian real*. Teorema următoare indică proprietăți utile ale corpului borelian real.

Teorema 3.2.4. *Dacă* $\mathcal{B}$ *este corpul borelian real, atunci:*

(*a*) $\mathrm{R} = (-\infty, +\infty) \in \mathcal{B}$;

(*b*) *Pentru orice* $\mathrm{a, b} \in \mathrm{R}$, *cu* $\mathrm{a} < \mathrm{b}$, *avem* $[\mathrm{a}, \mathrm{b}) \in \mathcal{B}$, $[\mathrm{a}, \mathrm{b}] \in \mathcal{B}$ *și* $(\mathrm{a}, \mathrm{b}) \in \mathcal{B}$;

(*c*) *Orice mulțime deschisă în* R *și orice mulțime închisă în* R *aparțin lui* $\mathcal{B}$.

Demonstrație. (*a*) Dacă (a_n) este un șir crescător de numere reale pentru care $a_n \to +\infty$, rezultă că

$$R = \bigcup_{n \in N} (-\infty, a_n) \in \mathfrak{B},$$

deoarece $(-\infty, a_n) \in \mathfrak{B}$.

(*b*). Avem $[a, b) = (-\infty, b) - (-\infty, a)$ și deci $[a, b) \in \mathfrak{B}$. Deoarece

$$[a, b] = \bigcap_{n \in N} [a, b + 1/n),$$

rezultă $[a, b] \in \mathfrak{B}$. În sfîrșit, deoarece $(a, b) = (-\infty, b) - (-\infty, a]$, rezultă $(a, b) \in \mathfrak{B}$.

(*c*). Deoarece o mulțime deschisă în R este reuniunea unui șir de intervale deschise din R, rezultă că ea este inclusă în $\mathfrak{B}$. Deoarece o mulțime închisă este complementara unei mulțimi deschise, rezultă că ea este de asemenea inclusă în $\mathfrak{B}$.

Analog, în cazul cînd E este spațiul real n-dimensional R^n, corpul borelian generat de mulțimile deschise din R^n este numit *corpul borelian n-dimensional real* și este notat prin $\mathfrak{B}^n$. Acest corp borelian va conține de asemenea toate tipurile de intervale din R^n, precum și mulțimile care se obțin prin operațiile $(CB_1) - (CB_2)$ din acestea. Perechea $(R^n, \mathfrak{B}^n)$ va fi numită *cîmpul borelian real n-dimensional*.

3.3. DEFINIȚIA CLASICĂ A PROBABILITĂȚII

Alături de noțiunea de eveniment incert introdusă mai sus, probabilitatea este un alt concept fundamental al teoriei probabilităților. Există mai multe definiții ale acestei noțiuni. În cele ce urmează ne vom ocupa de așa-numita *definiție clasică* a probabilității. Această definiție reduce conceptul de probabilitate la noțiunea de *egal probabilitate* (egal posibilitate), ultima noțiune fiind considerată noțiune primară, care nu se definește formal.

Pentru a fi mai preciși, să considerăm o desfacere finită $\mathfrak{D} = \{E_1, \ldots, E_n\}$ a evenimentului sigur E. Vom presupune că

aceste evenimente sînt egal probabile (egal posibile), înțelegînd prin aceasta că există motive puternice care ne permit să considerăm că nici unul dintre aceste evenimente nu are o șansă mai mare de producere față de celelalte evenimente din desfacere.

Definiția clasică a probabilității este dată pentru evenimentele care aparțin corpului borelian finit $\mathfrak{B}(\mathfrak{D})$ generat de desfacerea $\mathfrak{D}$; corpul borelian $\mathfrak{B}(\mathfrak{D})$ va fi notat și prin $\mathfrak{B}(E_1, ..., E_n)$.

Definiția 3.3.1. *Dacă evenimentul* $A \in \mathfrak{B}(\mathfrak{D})$ *se desface în* m *dintre evenimentele* $E_1, ..., E_n$, *atunci probabilitatea sa* $P(A)$ *este definită de*

$$P(A) = m/n.$$

Din această definiție rezultă imediat că probabilitatea evenimentului sigur este egală cu 1, probabilitatea evenimentului imposibil este egală cu 0, iar probabilitatea unui eveniment din corpul $\mathfrak{B}(E_1, ..., E_n)$ este un număr cuprins între 0 și 1.

În aplicațiile practice, evenimentele $E_1, ..., E_n$ reprezintă rezultatele posibile ale experimentului, care mai sînt numite și *cazuri*. Rezultatele posibile în care se desface un eveniment $A \in \mathfrak{B}(E_1, ..., E_n)$ sînt numite și *cazuri favorabile producerii* evenimentului A. Cu această terminologie, probabilitatea lui A este raportul dintre numărul cazurilor favorabile producerii lui A și numărul tuturor cazurilor egal probabile ale experimentului.

Exemple. 1. Să considerăm experimentul care constă în aruncarea unei monezi. Notăm prin E_1 și E_2 evenimentele care constau în apariția stemei, respectiv a valorii. Este clar că E_1 și E_2 formează o desfacere a evenimentului sigur E. Dacă nu avem motive să considerăm că una dintre fețe este mai favorizată decît cealaltă (moneda este deci perfectă, ideală), atunci rezultă că evenimentele E_1 și E_2 sînt egal probabile. Corpul de evenimente generat de această desfacere este

$$\mathfrak{B}(E_1, E_2) = \{\varnothing, E_1, E_2, E\},$$

de unde rezultă imediat că avem

$$P(E_1) = P(E_2) = 1/2.$$

2. Să considerăm experimentul care constă în aruncarea unui zar și să notăm prin $E_1, ..., E_6$ evenimentele care constau

respectiv în apariția fețelor 1, ..., 6 ale zarului. Presupunem că aceste evenimente sînt egal probabile (zarul este perfect, ideal).

Este clar că $\{E_1, ..., E_6\}$ formează o desfacere a evenimentului sigur E. Corpul borelian finit $\mathfrak{B}(E_1, ..., E_6)$ generat de această desfacere are $2^6 = 64$ evenimente. Este ușor de văzut că

$$P(E_1) = ... = P(E_6) = 1/6.$$

Dacă $A = E_1 \cup E_3 \cup E_5$ este evenimentul care constă din apariția unei fețe impare la aruncarea zarului, atunci avem $P(A) = 3/6$, numărul cazurilor favorabile producerii lui A fiind egal cu 3.

3. Să considerăm o urnă cu 7 bile albe și 5 bile roșii. Experimentul constă în extragerea unei bile din această urnă. Fie $A_1, ..., A_7$ și $R_1, ..., R_5$ evenimentele care constau în apariția uneia dintre cele 7 bile albe sau a uneia din cele 5 bile roșii (presupunem că bilele au fost numerotate). Este clar că cele 12 evenimente $A_1, ..., A_7, R_1, ..., R_5$ formează o desfacere a evenimentului sigur, iar aceste evenimente se pot considera ca fiind egal probabile. Evenimentul A care constă în apariția unei bile albe este $A = A_1 \cup ... \cup A_7$, iar evenimentul care constă în apariția unei bile roșii este $R = R_1 \cup ... \cup R_5$. Deoarece numărul cazurilor favorabile producerii lui A este 7, iar numărul cazurilor favorabile producerii lui R este 5, rezultă că avem $P(A) = 7/12$ și $P(R) = 5/12$.

Este clar că probabilitatea P poate fi considerată ca fiind o funcție definită pe corpul borelian $\mathcal{K} = \mathfrak{B}(E_1, ..., E_n)$ generat de o desfacere $\mathcal{D} = \{E_1, ..., E_n\}$ a evenimentului sigur E. Proprietățile acestei funcții sînt precizate în teorema următoare.

Teorema 3.3.1. *Cu definițiile de mai sus, dacă* P *este o probabilitate pe* $\mathcal{K} = \mathfrak{B}(E_1, ..., E_n)$, *atunci:*

(a) $P(\varnothing) = 1$, $P(E) = 1$ *și* $P(A) \in [0, 1]$ *pentru orice* $A \in \mathcal{K}$;

(b) Dacă $A_1, ..., A_k$ *este o desfacere a evenimentului* $A \in \mathcal{K}$, *atunci* $P(A) = P(A_1) + ... + P(A_k)$;

(c) Dacă $A \in \mathcal{K}$, *atunci* $P(A^c) = 1 - P(A)$;

(d) Dacă $A, B \in \mathcal{K}$ *și* $A \subset B$, *atunci* $P(A) \leqslant P(B)$.

Demonstrație. *(a)*. Este evidentă, conform definiției.

(b). Vom considera $k = 2$, cazul general rezultînd prin inducție după k. Fie m_1 numărul cazurilor favorabile lui A_1 și m_2 numărul cazurilor favorabile lui A_2. Deoarece $A_1 \cap A_2 = \varnothing$,

rezultă că numărul cazurilor favorabile lui $A = A_1 \cup A_2$ este $m = m_1 + m_2$. Rezultă atunci

$$P(A) = m/n = (m_1 + m_2)/n = m_1/n + m_2/n = P(A_1) + P(A_2).$$

(c). Rezultă imediat din (a) și (b), observînd că $A \cup A^c = E$, $A \cap A^c = \varnothing$.

(d). Deoarece $A \subset B$, rezultă că putem scrie

$$B = A \cup (A^c \cap B)$$

și deci

$$P(B) = P(A) + P(A^c \cap B) \geqslant P(A).$$

Așa cum am văzut și din exemplele date mai sus, definiția clasică a probabilității este deosebit de utilă în probleme simple, de natură combinatorială, care apar în practică. Este totuși necesar să subliniem faptul că această definiție este limitată numai la cazul experimentelor cu un număr finit de rezultate posibile. Mai mult, definiția presupune că aceste rezultate sînt egal probabile (posibile). Trecînd peste viciul de logică evident în această definiție, trebuie să remarcăm că, în foarte multe probleme practice, rezultatele posibile ale experimentului nu sînt egal probabile, ceea ce face evident inaplicabilă această definiție.

3.4. DEFINIȚIA AXIOMATICĂ A PROBABILITĂȚII

Aplicațiile numeroase și foarte importante pe care conceptele și rezultatele teoriei probabilităților le-au avut în diverse domenii științifice sau tehnice au impus la începutul secolului nostru așezarea acestei teorii pe baze matematice solide. Deși s-au încercat mai multe moduri de construcție axiomatică a teoriei probabilităților, cea care s-a impus în mod evident este axiomatizarea dată de Kolmogorov în anul 1933. În această axiomatizare se pornește de la o mulțime E de *evenimente elementare*, care sînt asimilate în practică cu mulțimea rezultatelor posibile ale unui experiment. Al doilea element al acestei axiomatici este un corp borelian de evenimente $\mathfrak{K}$ pe E. În sfîrșit, ultimul element al acestei axiomatici este o funcție P

definită pe corpul borelian $\mathcal{K}$ și avînd valori reale, care verifică proprietățile precizate în definiția următoare.

Definiția 3.4.1. *Funcția* $P : \mathcal{K} \to R$ *este numită probabilitate pe cîmpul borelian de evenimente* $(E, \mathcal{K})$ *dacă îndeplinește condițiile următoare:*

(P_1) *Pentru orice* $A \in \mathcal{K}$, $P(A) \geqslant 0$;

(P_2) $P(E) = 1$;

(P_3) *Pentru orice familie cel mult numărabilă de evenimente* $(A_i)_{i \in I} \in \mathcal{K}$, *incompatibile două cîte două,* $P\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} P(A_i)$.

Dacă $(E, \mathcal{K})$ este un cîmp borelian de evenimente și P este o probabilitate pe acest cîmp, atunci tripletul $(E, \mathcal{K}, P)$ este numit *cîmp borelian de probabilitate* sau *spațiu probabilist*.

Prima axiomă (P_1) cere ca probabilitatea oricărui eveniment să fie un număr pozitiv sau zero. A doua axiomă cere ca probabilitatea evenimentului sigur să fie egală cu 1. În sfîrșit, a treia axiomă cere ca probabilitatea reuniunii evenimentelor oricărei desfaceri să fie egală cu suma probabilităților evenimentelor care o compun.

Dacă $(E, \mathcal{K})$ este numai cîmp de probabilitate, adică dacă $\mathcal{K}$ este numai corp pe E, putem defini o funcție $P : \mathcal{K} \to R$ care să verifice axiomele (P_1), (P_2) și (P_3'), unde

(P_3') *Pentru orice evenimente incompatibile* $A, B \in \mathcal{K}$, *avem* $P(A \cup B) = P(A) + P(B)$.

Funcția P definită în acest mod este numită *probabilitate simplu aditivă*, pentru a o deosebi de funcția P care verifică axiomele $(P_1) - (P_3)$ și care este o *probabilitate complet (sigma) aditivă*. Evident, o probabilitate complet aditivă este și simplu aditivă; deoarece interesul major este evident acordat probabilității complet aditive, nu vom mai reveni asupra probabilității simplu aditive.

Teorema 3.4.1. *Dacă* $(E, \mathcal{K}, P)$ *este un cîmp borelian de probabilitate, atunci:*

(a) *Dacă* $A \subset B$ *cu* $A, B \in \mathcal{K}$, *atunci* $P(B - A) = P(B) - P(A)$;

(b) *Dacă* $A \subset B$ *cu* $A, B \in \mathcal{K}$, *atunci* $P(A) \leqslant P(B)$;

(c) *Dacă* $A \in \mathcal{K}$, *atunci* $P(A^c) = 1 - P(A)$;

(d) $P(\varnothing) = 1$;

(e) *Dacă* $A \in \mathcal{K}$, *atunci* $0 \leqslant P(A) \leqslant 1$;

(f) *Dacă* $A, B \in \mathcal{K}$, *atunci* $P(A \cup B) = P(A) + P(B) - P(A \cap B)$;

(*g*) *Dacă* $(A_i)_{i\in I}\subset\mathfrak{K}$ *este o familie de evenimente cel mult numărabilă, atunci* $P\left(\bigcup_{i\in I} A_i\right)\leqslant\sum_{i\in I} P(A_i)$

și

$$P\left(\bigcap_{i\in I} A_i\right)\geqslant 1-\sum_{i\in I} P(A_i^c),$$

ultima inegalitate fiind cunoscută sub numele de inegalitatea lui Boole.

Demonstrație. (*a*). Pentru oricare două evenimente A și B din $\mathfrak{K}$, avem

$$B=(B-A)\cup(A\cap B)$$

cu

$$(B-A)\cap(A\cap B)=\varnothing.$$

Conform axiomei (P_3), rezultă că

$$P(B)=P(B-A)+P(A\cap B),$$

de unde se obține imediat relația căutată.

(*b*). Afirmația se obține din (*a*) și din faptul că $P(B-A)\geqslant 0$.

(*c*). Deoarece $E=A\cup A^c$ și $A\cap A^c=\varnothing$, rezultă că $P(E)=P(A)+P(A^c)$, de unde obținem relația dorită.

(*d*). Deoarece $\varnothing=E^c$, rezultatul urmează din cel precedent.

(*e*). Deoarece $\varnothing\subset A\subset E$ pentru orice $A\in\mathfrak{K}$, inegalitatea căutată rezultă imediat din (*b*).

(*f*). Pentru orice evenimente A, $B\in\mathfrak{K}$, avem

$$A\cup B=A\cup(B-A\cap B)$$

cu

$$A\cap(B-A\cap B)=\varnothing.$$

Rezultă atunci

$$P(A\cup B)=P(A)+P(B-A\cap B).$$

Din faptul că $A\cap B\subset B$ rezultă

$$P(B-A\cap B)=P(B)-P(A\cap B),$$

conform cu (*a*). Relația căutată se deduce imediat din ultimele două relații obținute mai sus.

Scanned with OKEN Scanner

(g). Pentru fiecare $i \in I$ definim evenimentul

$$B_i = A_i - A_1 \cup \ldots \cup A_{i-1}.$$

Este evident că $B_i \subset A_i$ pentru orice $i \in I$ și că

$$\bigcup_{i \in I} B_i = \bigcup_{i \in I} A_i.$$

Se poate acum observa că, pentru $i \neq j$, avem $B_i \cap B_j = \varnothing$. Conform axiomei (P_3) putem scrie atunci

$$\sum_{i \in I} P(B_i) = P\left(\bigcup_{i \in I} B_i\right).$$

Deoarece $B_i \subset A_i$, rezultă imediat

$$\sum_{i \in I} P(B_i) \leqslant \sum_{i \in I} P(A_i),$$

conform cu proprietatea (b).

Din ultimele două relații, obținem

$$P\left(\bigcup_{i \in I} A_i\right) = P\left(\bigcup_{i \in I} B_i\right) \leqslant \sum_{i \in I} P(A_i),$$

de unde rezultă prima inegalitate.

Pentru demonstrarea inegalității lui Boole, să observăm că

$$\bigcap_{i \in I} A_i = \left(\bigcup_{i \in I} A_i^c\right)^c = E - \bigcup_{i \in I} A_i^c$$

și deci

$$P\left(\bigcap_{i \in I} A_i\right) = 1 - P\left(\bigcup_{i \in I} A_i^c\right).$$

Inegalitatea lui Boole se obține din inegalitatea precedentă și ultima relație obținută.

În afara proprietăților expuse în teorema precedentă, probabilitatea mai are o proprietate importantă, numită *proprietatea de continuitate*, care, așa cum vom vedea, poate înlocui axioma de aditivitate completă în definiția probabilității.

Scanned with OKEN Scanner

Teorema 3.4.2. *Dacă* $(E, \mathcal{K}, P)$ *este un cîmp borelian de probabilitate, atunci:*

(a) Dacă $(A_n)_{n\in N} \subset \mathcal{K}$ *este un șir monoton descrescător de evenimente (adică* $A_n \supset A_{n+1}$ *pentru orice* $n \in N$*), atunci*

$$\lim_n P(A_n) = P\left(\bigcap_{n\in N} A_n\right);$$

(b) Dacă $(A_n)_{n\in N} \subset \mathcal{K}$ *este un șir monoton crescător de evenimente (adică* $A_n \subset A_{n+1}$ *pentru orice* $n \in N$*), atunci*

$$\lim_n P(A_n) = P\left(\bigcup_{n\in N} A_n\right).$$

Demonstrație. (*a*). Fie $A = \bigcap_{n\in N} A_n$ și fie $B_n = A_n - A$. Este clar că $\bigcap_{n\in N} B_n = \varnothing$ și că $(B_n)_{n\in N}$ este un șir descrescător de evenimente din $\mathcal{K}$. Deoarece

$$P(B_n) = P(A_n - A) = P(A_n) - P(A),$$

rezultă că pentru demonstrarea afirmației este suficient să arătăm că

$$\lim_n P(B_n) = 0.$$

Pentru aceasta să observăm că evenimentul B_n poate fi scris sub forma

$$B_n = (B_n - B_{n+1}) \cup (B_{n+1} - B_{n+2}) \cup \ldots$$

și, în virtutea axiomei (P_3), rezultă

$$P(B_n) = \sum_{m=n}^{\infty} P(B_m - B_{m+1}).$$

În particular, pentru $n = 0$, avem

$$P(B_0) = \sum_{n\in N} P(B_n - B_{n+1}).$$

Din ultimele două relații rezultă că $P(B_n)$ reprezintă restul seriei convergente scrisă în ultima relație. Urmează deci că $P(B_n) \to 0$ și deci relația căutată este demonstrată.

Scanned with OKEN Scanner

(b). Să considerăm șirul de evenimente cu termenul general $C_n = A_n^c$. Rezultă că $(C_n)_{n\in N}$ este un șir descrescător de evenimente din $\mathcal{K}$. Din (a) rezultă că

$$\lim_n P(C_n) = P\left(\bigcap_{n\in N} C_n\right),$$

adică

$$\lim_n (1 - P(A_n)) = 1 - P\left(\bigcup_{n\in N} A_n\right),$$

de unde se deduce imediat relația căutată.

Teorema 3.4.3. (Kolmogorov). *Condiția necesară și suficientă ca probabilitatea simplu aditivă* P *pe cîmpul borelian de evenimente* (E, $\mathcal{K}$) *să fie complet aditivă este ca, pentru orice șir descrescător de evenimente* $(A_n)_{n\in N}\subset\mathcal{K}$ *pentru care* $\bigcap_{n\in N} A_n = \varnothing$, *să avem*

$$\lim_n P(A_n) = 0.$$

Demonstrație. Necesitatea condiției rezultă din teorema precedentă.

Pentru demonstrarea suficienței, să considerăm șirul de evenimente $(B_n)_{n\in N}\subset\mathcal{K}$, incompatibile două cîte două. Fie

$$B = \bigcup_{n\in N} B_n,\quad A_n = B - \bigcup_{m\leqslant n} B_m.$$

Este clar că șirul de evenimente $(A_n)_{n\in N}\subset\mathcal{K}$ astfel definit este descrescător și

$$\bigcap_{n\in N} A_n = \bigcap_{n\in N}\left(B - \bigcup_{m\leqslant n} B_m\right) = \varnothing.$$

Rezultă atunci că șirul (A_n) are proprietățile cerute în teoremă și deci avem

$$\lim_n P(A_n) = 0,$$

adică

$$\lim_n \left(P(B) - P\left(\bigcup_{m\leqslant n} B_m\right)\right) = 0.$$

Scanned with OKEN Scanner

Din aditivitatea simplă a lui P și din ultima relație se obține

$$\lim_n \sum_{m \leqslant n} P(B_m) = P(B),$$

adică

$$\sum_{n \in N} P(B_n) = P\left(\bigcup_{n \in N} B_n\right)$$

și deci P este complet aditivă.

Definiția 3.4.2. *Cîmpul borelian de probabilitate* (E, $\mathcal{K}$, P) *este numit complet dacă pentru orice* $A \in \mathcal{K}$ *de probabilitate nulă și pentru orice* $B \subset A$ *avem* $B \in \mathcal{K}$.

Cu alte cuvinte, un corp borelian este complet dacă el conține toate evenimentele care implică evenimentele sale de probabilitate nulă. Este clar că, dacă $(E, \mathcal{K}, P)$ este complet și $A \in \mathcal{K}$ are probabilitatea egală cu 1, atunci orice eveniment B implicat de A este de asemenea în $\mathcal{K}$. Într-adevăr, deoarece $A^c \in \mathcal{K}$ și $P(A^c) = 0$, rezultă că $B^c \subset A^c$ este de asemenea în $\mathcal{K}$ și deci $B \in \mathcal{K}$.

Se poate arăta că orice cîmp borelian de probabilitate poate fi completat. Mai precis, dacă $(E, \mathcal{K}, P)$ este un cîmp borelian de probabilitate, atunci se poate arăta că există un cîmp borelian de probabilitate $(E, \mathcal{K}^+, P^+)$ complet cu proprietatea că $\mathcal{K} \subset \mathcal{K}^+$, iar restricția funcției P^+ la $\mathcal{K}$ coincide cu P. În cele ce urmează vom presupune cîmpurile de probabilitate considerate ca fiind complete.

În încheiere, să observăm că sistemul de axiome al lui Kolmogorov este consistent. Pentru aceasta este suficient să furnizăm exemple de cîmpuri boreliene de probabilitate.

Exemplul 1. Să considerăm mulțimea finită $E = \{\omega_1, \ldots, \omega_n\}$ și corpul borelian total $\mathcal{K}_t = \mathfrak{P}(E)$. Să considerăm n numere reale nenegative $p_1, \ldots, p_n$, a căror sumă este egală cu 1. Pe cîmpul borelian de evenimente $(E, \mathcal{K}_t)$ definim funcția P în modul următor: pentru orice eveniment $A = \{\omega_{i_1}, \ldots, \omega_{i_k}\}$ punem

$$P(A) = \sum_{j=1}^{k} p_{i_j}.$$

În particular, $P(\{\omega_i\}) = p_i$, $1 \leqslant i \leqslant n$.

Se poate vedea ușor că funcția P astfel definită este o probabilitate pe cîmpul borelian de evenimente $(E, \mathcal{K}_t)$ și deci $(E, \mathcal{K}_t, P)$ este un cîmp borelian de probabilitate.

Exemplul 2. Să considerăm mulțimea numărabilă $E = (\omega_i)_{i \in I}$ și corpul borelian total $\mathcal{K}_t = \mathcal{P}(E)$. Să considerăm mulțimea $(p_i)_{i \in I}$ de numere reale nenegative cu proprietatea că $\sum_{i \in I} p_i = 1$. Pe cîmpul borelian de evenimente $(E, \mathcal{K}_t)$ definim funcția $P : \mathcal{K}_t \to R$ dată prin

$$P(A) = \sum_{i \in J} p_i$$

pentru orice eveniment $A = (\omega_i)_{i \in J}$, unde $J \subset I$.

Se poate arăta ușor că funcția P astfel definită este o probabilitate complet aditivă pe $\mathcal{K}_t$ și deci $(E, \mathcal{K}_t, P)$ este un cîmp borelian de probabilitate.

Să mai observăm că sistemul de axiome al lui Kolmogorov nu definește complet probabilitățile evenimentelor. Într-adevăr, dacă ne referim la primul exemplu considerat mai sus, este ușor de văzut că numerele $p_1, \ldots, p_n$ sînt supuse la condițiile următoare

$$p_i \geqslant 0, \quad 1 \leqslant i \leqslant n,$$

$$p_1 + \ldots + p_n = 1,$$

în rest fiind arbitrare. Definiția clasică a probabilității, care se obține considerînd că probabilitățile p_i, $1 \leqslant i \leqslant n$, sînt egale, adică $p_i = 1/n$, $1 \leqslant i \leqslant n$, ne dă valorile precise ale acestor probabilități, părînd la prima vedere superioară.

Faptul că sistemul de axiome nu definește complet probabilitățile evenimentelor nu constituie însă o slăbiciune a acestui sistem. Dimpotrivă, „incompletitudinea" sistemului este esențială în aplicațiile practice, în care trebuie considerate familii de evenimente identice, dar avînd probabilități diferite. De exemplu, aruncarea unei monede perfecte ne conduce la considerarea probabilităților egale, pe cînd aruncarea unei monede cu defecte ne conduce la considerarea de probabilități inegale asociate celor două fețe.

Scanned with OKEN Scanner

3.5. PROBABILITATE CONDIȚIONATĂ

Să revenim pentru scurt timp la definiția clasică a probabilității. În anumite situații sîntem conduși la calculul probabilității unui eveniment A în ipoteza că un alt eveniment B de probabilitate nenulă s-a produs. O astfel de probabilitate va fi numită *condiționată* și va fi notată prin $P(A \mid B)$ sau $P_B(A)$.

Pentru a stabili o formulă pentru calculul acestei probabilități, să notăm cu n numărul rezultatelor posibile ale unui anumit experiment. Dacă numărul rezultatelor favorabile producerii evenimentului B este r $(r \leqslant n)$, atunci $P(B) = r/n$. Dacă numărul rezultatelor favorabile producerii evenimentului $A \cap B$ este s (evident, $s \leqslant n$ și $s \leqslant r$), atunci $P(A \cap B) = s/n$.

Să arătăm acum că $P(A \mid B) = s/r$. Într-adevăr, numărul total al cazurilor posibile este în acest caz numărul cazurilor favorabile producerii lui B. Dacă B s-a produs, atunci numărul cazurilor favorabile lui A este același cu numărul cazurilor favorabile lui $A \cap B$, de unde se deduce rezultatul dorit.

Un raționament analog ne conduce la concluzia că $P(B \mid A) = r/n$.

Din relațiile obținute mai sus se deduc fără dificultate formulele căutate:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \tag{3.5.1}$$

și

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}. \tag{3.5.2}$$

Este clar că formula (3.5.1) poate să nu aibă sens pentru $P(B)=0$, iar formula (3.5.2) poate să nu aibă sens pentru $P(A) = 0$.

Din formulele precedente se obține relația

$$P(A \cap B) = P(A)\, P(B|A) = P(B)\, P(A|B), \tag{3.5.3}$$

numită și *regula de înmulțire a probabilităților* (vezi teorema 3.5.1 pentru rezultatul general). Această regulă se menține și în cazul cînd unul dintre evenimentele A sau B este de probabilitate nulă.

Exemplu. Să considerăm o urnă care conține 3 bile albe și 5 bile roșii. Se extrag două bile. Să calculăm probabilitatea ca ambele bile extrase să fie roșii. Se poate, evident, presupune că cele două bile se extrag pe rînd, prima bilă extrasă nefiind însă introdusă în urnă înainte de extragerea a doua. Să notăm prin R_1 evenimentul care constă în apariția unei bile roșii la prima extragere și prin R_2 evenimentul care constă în apariția unei bile roșii la extragerea a doua. Evenimentul care ne interesează este evident $R_1 \cap R_2$. Avem, mai întîi, $P(R_1) = 5/8$, iar apoi $P(R_2 \mid R_1) = 4/7$, deoarece, dacă prima bilă extrasă este roșie, atunci numărul bilelor roșii din urnă devine $5 - 1 = 4$, iar numărul total al bilelor devine $8 - 1 = 7$. Rezultă atunci

$$P(R_1 \cap R_2) = P(R_1)\, P(R_2 \mid R_1) = 5/14.$$

Așa cum am văzut mai sus, în cadrul definiției clasice a probabilității am reușit să demonstrăm formulele (3.5.1) și (3.5.2), care leagă noțiunea de probabilitate condiționată de conceptul de probabilitate introdus anterior. În cazul definiției axiomatice a lui Kolmogorov acest lucru nu mai este posibil, definiția probabilității condiționate dîndu-se pe baza formulei (3.5.1).

Pentru a fi mai precizi, să considerăm un cîmp borelian de probabilitate $(E, \mathcal{K}, P)$ și fie $B \in \mathcal{K}$ un eveniment de probabilitate nenulă. Pentru orice $A \in \mathcal{K}$ definim probabilitatea sa condiționată prin formula (3.5.1). Se poate vedea ușor că această formulă ne permite să definim pe $\mathcal{K}$ funcția cu valori reale $P(. \mid B)$, care asociază fiecărui $A \in \mathcal{K}$ numărul real $P(A \mid B)$. Vom arăta că această funcție este chiar o probabilitate complet aditivă pe corpul $\mathcal{K}$, adică verifică axiomele $(P_1) - (P_3)$.

Verificarea axiomei (P_1) este trivială. Pentru verificarea axiomei (P_2), să observăm că avem

$$P(E \mid B) = \frac{P(E \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1.$$

În sfîrșit, pentru verificarea axiomei (P_3), să considerăm familia cel mult numărabilă $(A_i)_{i \in I} \subset \mathcal{K}$ de evenimente incompa-

tibile două cîte două. Avem

$$P\left(\bigcup_{i\in I} A_i \mid B\right) = \frac{P\left(\bigcup_{i\in I}(B\cap A_i)\right)}{P(B)} = \frac{\sum_{i\in I} P(B\cap A_i)}{P(B)} =$$

$$= \sum_{i\in I} \frac{P(B\cap A_i)}{P(B)} = \sum_{i\in I} P(A_i \mid B),$$

ceea ce demonstrează că $(E, \mathcal{K}, P(.\mid B))$ este un cîmp borelian de probabilitate* și deci $P(.\mid B)$ are toate proprietățile pe care le-am demonstrat în § 3.4 pentru probabilitățile complet aditive. În plus, probabilitatea condiționată are și proprietățile precizate în teoremele următoare.

Teorema 3.5.1. *Fie* $(E, \mathcal{K}, P)$ *un cîmp borelian de probabilitate și fie* $A_1, \ldots, A_n$ *evenimente din* $\mathcal{K}$ *cu proprietatea că* $P(A_1\cap\ldots\cap A_{n-1})\neq 0$. *În aceste condiții, avem*

$$P(A_1\cap\ldots\cap A_n)=P(A_1)P(A_2\mid A_1)\ldots P(A_n\mid A_1\cap\ldots\cap A_{n-1}), \quad (3.5.4)$$

formulă cunoscută sub numele de **regula de înmulțire a probabilităților.**

Demonstrație. Formula (3.5.4) se demonstrează prin inducție după $n\geqslant 2$. Pentru $n=2$ formula este evidentă, reducîndu-se la definiție. Presupunînd-o adevărată pentru $n-1$, rezultă

$$P(A_1\cap\ldots\cap A_{n-1}\cap A_n)=P(A_1\cap\ldots\cap A_{n-1})P(A_n\mid A_1\cap\ldots\cap A_{n-1})=$$
$$=P(A_1)P(A_2\mid A_1)\ldots P(A_{n-1}\mid A_1\cap\ldots\cap A_{n-2})P(A_n\mid A_1\cap\ldots\cap A_{n-1}),$$

ceea ce arată că formula (3.5.4) este valabilă pentru orice n.

Teorema 3.5.2. *Fie* $(E, \mathcal{K}, P)$ *un cîmp borelian de probabilitate și fie* $(A_i)_{i\in I}\subset\mathcal{K}$ *o familie cel mult numărabilă de evenimente incompatibile două cîte două, avînd proprietatea că* $P(A_i)>0$ *pentru orice* $i\in I$ *și* $\bigcup_{i\in I} A_i = E$. *În aceste condiții, pentru orice* $A\in\mathcal{K}$, *avem*

$$P(A) = \sum_{i\in I} P(A_i)\,P(A\mid A_i), \quad (3.5.5)$$

relație cunoscută și sub numele de **formula probabilității totale.**

---

* De fapt, $P(.\mid B)$ este o probabilitate complet aditivă pe corpul borelian $\mathcal{K}_B = \{A\cap B : A\in\mathcal{K}\}$.

Demonstrație. Avem, evident,

$$A = A \cap E = A \cap \left(\bigcup_{i \in I} A_i \right) = \bigcup_{i \in I} (A \cap A_i).$$

Deoarece evenimentele din familia $(A \cap A_i)_{i \in I}$ sînt evident incompatibile două cîte două, rezultă că putem scrie

$$P(A) = P\left(\bigcup_{i \in I} (A \cap A_i) \right) = \sum_{i \in I} P(A \cap A_i) = \sum_{i \in I} P(A_i)\, P(A \mid A_i),$$

adică tocmai ce trebuia demonstrat.

Teorema 3.5.3 *(Formula lui Bayes). Fie* $(\mathrm{E}, \mathfrak{K}, \mathrm{P})$ *un cîmp borelian de probabilitate și* $(\mathrm{A_i})_{i \in \mathrm{I}} \subset \mathfrak{K}$ *o familie cel mult numărabilă de evenimente incompatibile două cîte două cu proprietatea că* $\mathrm{P(A_i)} > 0$ *pentru orice* $i \in \mathrm{I}$ *și* $\bigcup_{i \in \mathrm{I}} \mathrm{A_i} = \mathrm{E}$.

În aceste condiții, pentru orice $\mathrm{A} \in \mathfrak{K}$ *pentru care avem* $\mathrm{P(A)} > 0$, *este adevărată formula lui Bayes:*

$$\mathrm{P(A_i \mid A)} = \frac{\mathrm{P(A_i)\, P(A \mid A_i)}}{\sum_{i \in \mathrm{I}} \mathrm{P(A_i)\, P(A \mid A_i)}}. \tag{3.5.6}$$

Demonstrație. Avem

$$P(A_i \mid A) = \frac{P(A \cap A_i)}{P(A)} = \frac{P(A_i)\, P(A \mid A_i)}{P(A)}.$$

Ținînd seama de (3.5.5), formula (3.5.6) se deduce imediat din relația precedentă.

Observație. Formula lui Bayes este de o importanță fundamentală într-o serie de aplicații ale calculului probabilităților. Procedeul general de aplicare al acestei formule la probleme de evaluare a probabilităților unui sistem complet de evenimente $(A_i)_{i \in I}$ este următorul. Se cunosc *probabilitățile inițiale* sau apriori $(P(A_i))_{i \in I}$ pentru acest sistem de evenimente. Se face un experiment în urma căruia se produce un eveniment A. Presupunînd că probabilitățile condiționate $(P(A \mid A_i))_{i \in I}$ se pot calcula, formula lui Bayes ne permite să evaluăm probabilitățile sistemului complet de evenimente $(A_i)_{i \in I}$ pe baza

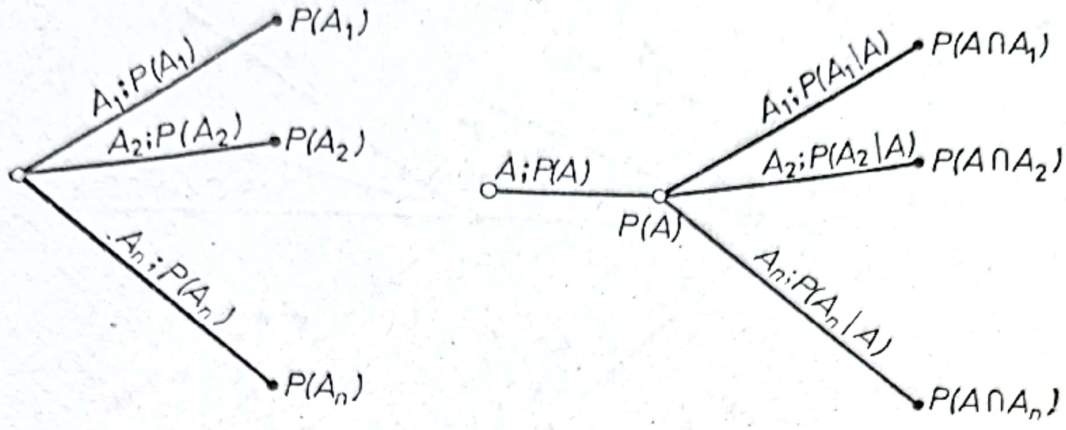


Fig. 3.5.1

cunoașterii rezultatului experimentului efectuat; noile probabilități vor fi desigur $(P(A_i|A))_{i\in I}$ și vor fi numite *probabilități aposteriori* sau *finale*.

În aplicațiile practice este comod să prezentăm un sistem complet de evenimente $(A_1, ..., A_n) \subset \mathcal{K}$ printr-un *arbore* (vezi fig. 3.5.1 (a)). Fiecare muchie corespunde cîte unui eveniment și se asociază de asemenea cu probabilitatea acelui eveniment. Dacă A este un eveniment de probabilitate nenulă, atunci putem de asemenea să asociem sistemului complet de evenimente $(A_1, ..., A_n) \subset \mathcal{K}$ un arbore (vezi fig. 3.5.1 (b)), utilizînd probabilitatea condiționată $P(. \mid A)$ în loc de probabilitatea $P(.)$.

Se poate observa că fiecare drum dintr-un astfel de arbore poate fi asociat cu un eveniment. De exemplu, drumul format în arborele din fig. 3.5.1 (a) de singura muchie A_1 este asociat evenimentului A_1, iar drumul format în arborele din fig. 3.5.1 (b) de muchiile A și A_1 este asociat evenimentului $A \cap A_1$. Probabilitățile asociate evenimentelor corespunzătoare acestor drumuri se înscriu în nodurile finale ale drumurilor. În acest mod se asociază fiecărui nod al grafului, cu excepția nodului inițial, probabilitatea de a ajunge în acest nod, respectiv probabilitatea asociată drumului care se termină în acest nod (vezi și fig. 3.5.1 (a) și (b)).

În multe probleme practice sîntem conduși la considerarea unei succesiuni de sisteme complete de evenimente; rezultatele posibile ale unei succesiuni de experimente oferă un astfel de exemplu. Unei succesiuni de sisteme complete de evenimente i se poate asocia un arbore în modul următor. Dacă

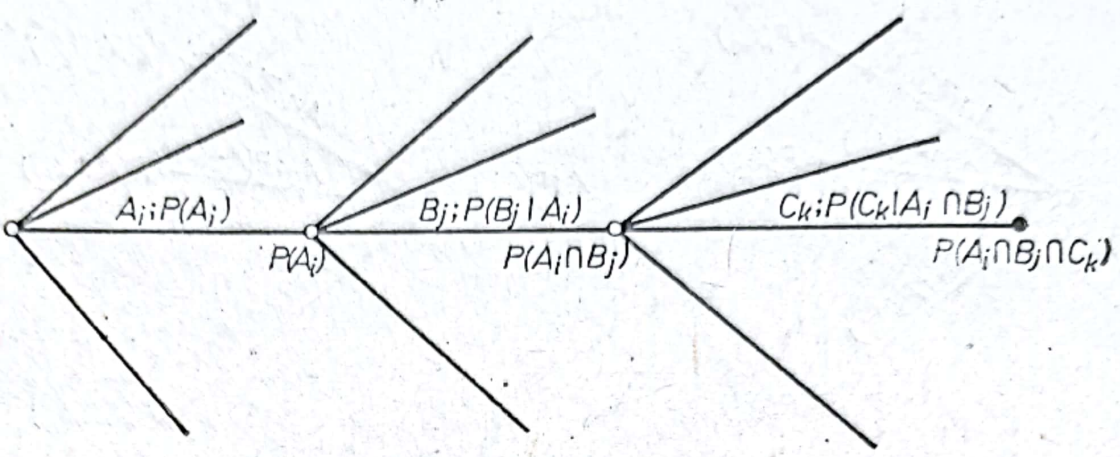


Fig. 3.5.2

de exemplu, avem trei sisteme complete de evenimente $(A_1, \ldots, A_m)$, $(B_1, \ldots, B_n)$ și $(C_1, \ldots, C_p)$, atunci arborele asociat este prezentat în fig. 3.5.2. Observațiile anterioare privind drumurile din acest arbore sînt evident valabile și în acest caz.

Exemplu. Pentru a ilustra modul de aplicare al teoremei lui Bayes, să considerăm următoarea problemă simplă. Avem 100 urne cu următoarele compoziții: 30 urne cu 6 bile albe și 4 bile roșii (compoziția A_1), 50 urne cu 2 bile albe și 3 bile roșii (compoziția A_2) și 20 urne cu 8 bile albe și 2 bile roșii (compoziția A_3). Urnele sînt identice ca aspect exterior și în același timp opace, astfel încît nu se poate vedea compoziția lor. Se extrage o bilă dintr-o urnă aleasă la întîmplare printre cele 100 urne disponibile. Se cer probabilitățile aposteriori ca extragerea să fi avut loc dintr-o urnă de compoziție A_i $(i = 1, 2, 3)$, dacă bila extrasă este roșie (evenimentul R), respectiv albă (evenimentul A).

Probabilitățile apriori ca extracția să aibă loc dintr-o urnă de compoziție A_1, A_2, A_3 sînt respectiv:

$$P(A_1) = 30/100; \quad P(A_2) = 50/100; \quad P(A_3) = 20/100.$$

Probabilitățile condiționate $P(R \mid A_i)$, $i = 1, 2, 3$, sînt:

$$P(R \mid A_1) = 2/5; \quad P(R \mid A_2) = 3/5; \quad P(R \mid A_3) = 1/5.$$

Analog, probabilitățile condiționate $P(A \mid A_i)$, $i = 1, 2, 3$, sînt:

$$P(A \mid A_1) = 3/5; \quad P(A \mid A_2) = 2/5; \quad P(A \mid A_3) = 4/5.$$

Scanned with OKEN Scanner

Aplicînd formula probabilității totale, avem

$$P(R)=P(A_1)\,P(R|A_1)+P(A_2)\,P(R|A_2)+P(A_3)\,P(R|A_3)=23/50.$$

Rezultă atunci că avem

$$P(A) = 1 - P(R) = 27/50.$$

Aplicînd formula lui Bayes, avem:

$$P(A_1 \mid R)= \frac{P(A_1)\,P(R \mid A_1)}{P(R)} = 6/23;$$

$$P(A_2 \mid R)= \frac{P(A_2)\,P(R \mid A_2)}{P(R)} = 15/23$$

și

$$P(A_3 \mid R)= \frac{P(A_3)\,P(R \mid A_3)}{P(R)} = 2/23.$$

Analog, obținem

$$P(A_1 \mid A) = 9/27; \quad P(A_2 \mid A) = 10/27; \quad P(A_3 \mid A) = 8/27.$$

Modificarea valorilor probabilităților aposteriori (în funcție de rezultatul experimentului) față de probabilitățile apriori reflectă importanța experimentului, valoarea sa informativă. Astfel, de exemplu, apariția bilei roșii ne conduce la aprecierea că șansa ca extragerea să fi avut loc dintr-o urnă de compoziție A_2 crește de la 1/2 la 15/23.

Calculele de mai sus se pot face și cu ajutorul arborelui asociat acestui experiment (vezi fig. 3.5.3). În figura 3.5.3 (a) este reprezentat arborele asociat situației apriori (se utilizează probabilitățile apriori pentru A_1, A_2, A_3), iar în fig. 3.5.3 (b) avem arborele asociat situației aposteriori (se utilizează probabilitățile aposteriori pentru A_1, A_2, A_3).

Arborele asociat situației inițiale se completează imediat. El permite calculul simplu al probabilităților evenimentelor $A_i \cap R$ și $A_i \cap A$, $1 \leqslant i \leqslant 3$. Într-adevăr, avem:

$$P(A_1 \cap R) = P(A_1)\,P(R \mid A_1) = 6/50;$$

$$P(A_2 \cap R) = P(A_2)\,P(R \mid A_2) = 15/50;$$

$$P(A_3 \cap R) = P(A_3)\,P(R \mid A_3) = 2/50$$

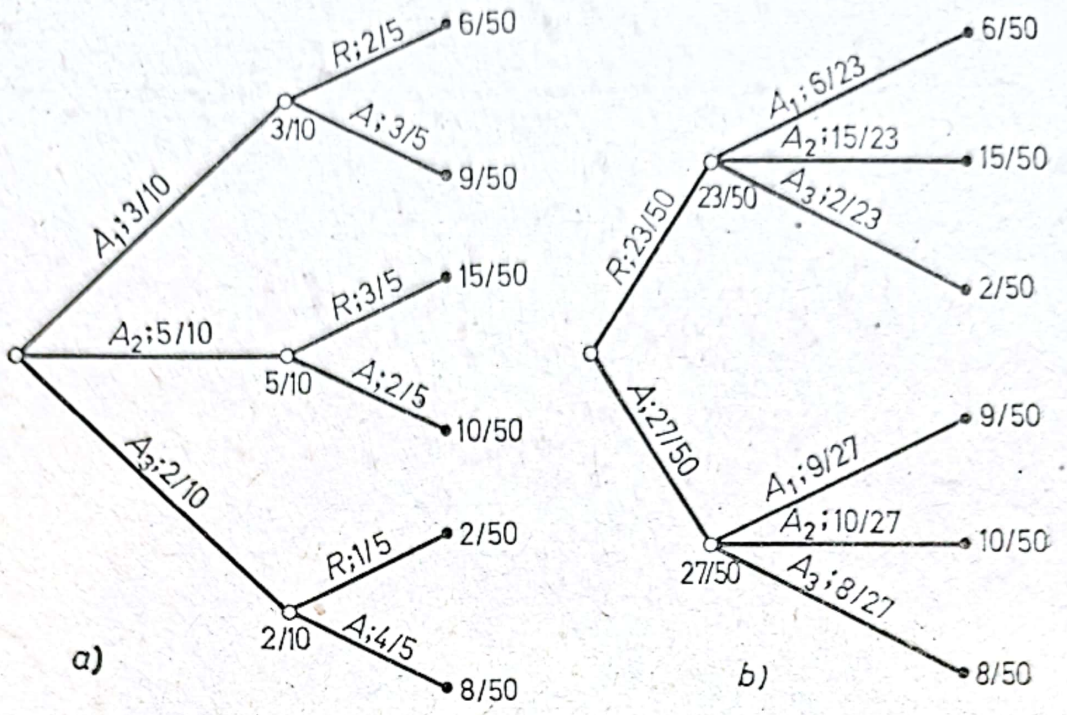


Fig. 3.5.3

și deci, pentru calculul acestor probabilități, este suficient să facem produsul probabilităților de pe drumurile (A_1, R), (A_2, R) și (A_3, R) ale arborelui asociat; rezultatele se trec în dreptul nodurilor finale ale drumurilor respective (vezi și fig. 3.5.2 (a)). Procedînd analog pentru drumurile (A_1, A), (A_2, A) și (A_3, A) se obțin numerele 9/50, 10/50 și 8/50, pe care le trecem de asemenea în dreptul nodurilor finale din arborele asociat situației apriori (vezi fig. 3.5.3 (a)). Probabilitățile asociate celorlalte noduri sînt evident $P(A_1)$, $P(A_2)$ și $P(A_3)$.

Să trecem la arborele asociat situației aposteriori (vezi fig. 3.5.3 (b)). Evident, drumurile (R, A_1) și (A_1, R) sînt ambele asociate evenimentului $A_1 \cap R$ și deci avem aceeași probabilitate la nodul final al drumului (R, A_1) din fig. 3.5.3 (b) ca și la nodul final al drumului (A_1, R) din fig. 3.5.3 (a). Cu alte cuvinte, probabilitățile asociate nodurilor finale din fig. 3.5.3 (b) sînt deja calculate și se trec direct pe această figură.

Putem acum calcula probabilitățile evenimentelor R și A. Avem

$$P(R) = P(R \cap A_1) + P(R \cap A_2) + P(R \cap A_3) = 23/50$$

și, analog,

$$P(A) = P(A \cap A_1) + P(A \cap A_2) + P(A \cap A_3) = 27/50,$$

probabilități pe care le trecem pe muchiile R și A din arborele asociat situației aposteriori, precum și în nodurile finale corespunzătoare acestor muchii.

Calculul probabilităților condiționate $P(A_i \mid R)$ și $P(A_i \mid A)$, $1 \leqslant i \leqslant 3$, care mai lipsesc din acest arbore, se face după formula probabilității condiționate. De exemplu, avem

$$P(A_1 \mid R) = P(A_1 \cap R) \,/\, P(R) = 6/23.$$

Cu alte cuvinte, probabilitatea asociată muchiei A_1 se calculează împărțind numărul 6/50 scris în nodul final prin numărul 23/50 scris în nodul inițial al acestei muchii. Analog, obținem numerele 15/23 și 2/23 pentru muchiile A_2 și A_3. Procedînd la fel pentru calculul probabilităților condiționate $P(A_1|A)$, $P(A_2|A)$ și $P(A_3|A)$, obținem numerele 9/27, 10/27 și 8/27 pe care le trecem de asemenea pe muchiile corespunzătoare.

3.6. EVENIMENTE DEPENDENTE ȘI EVENIMENTE INDEPENDENTE

Să considerăm un cîmp de probabilitate $(E, \mathfrak{K}, P)$ și evenimentele A, $B \in \mathfrak{K}$. Vom spune că evenimentele A și B sînt *independente* dacă

$$P(A \cap B) = P(A)\ P(B). \tag{3.6.1}$$

În caz contrar, evenimentele sînt numite *dependente*.

Dacă (A, B) formează un cuplu de evenimente independente, atunci evenimentele din cuplurile (A^c, B), (A, B^c) și (A^c, B^c) sînt de asemenea independente. Într-adevăr, pentru primul cuplu, avem:

$$P(A^c \cap B) = P((E - A) \cap B) = P(B - A \cap B) = P(B) - P(A \cap B) =$$
$$= P(B) - P(A)\ P(B) = P(B)\,(1 - P(A)) = P(A^c)\ P(B),$$

ceea ce demonstrează afirmația.

Noțiunea de independență a evenimentelor este de o importanță deosebită în teoria probabilităților. Pentru a determina în aplicațiile practice dacă evenimentele A și B sînt independente nu se folosește în general definiția formală de mai sus (adică verificarea relației (3.6.1)). În mod obișnuit se rezolvă problema intuitiv. De exemplu, apariția stemei la aruncarea unei monede este un eveniment independent de apariția stemei la aruncarea altei monede.

Vom trece acum la generalizarea noțiunii de independență pentru n evenimente $A_1, \ldots, A_n$ din corpul $\mathcal{K}$. Vom spune că evenimentele $A_1, \ldots, A_n$ sînt *independente m cîte m*, cu $m \leqslant n$, dacă pentru orice s ($s \leqslant m$) și pentru orice $1 \leqslant i_1 < \ldots < i_s \leqslant m$, avem

$$P(A_{i_1} \cap \ldots \cap A_{i_s}) = P(A_{i_1}) \ldots P(A_{i_s}). \tag{3.6.2}$$

Din această definiție rezultă că, dacă evenimentele $A_1, \ldots, A_n$ sînt independente m cîte m, atunci ele sînt independente și m' cîte m', dacă $m' \leqslant m$. Reciproca acestei afirmații este în general falsă. Pentru aceasta este suficient să dăm următorul exemplu, datorat lui S. N. Bernstein. Fie cîmpul de probabilitate $(E, \mathcal{K}_t, P)$, unde $E = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, $\mathcal{K}_t = \mathfrak{L}(E)$, iar P este definită de $P(\{\omega_i\}) = 1/4$ pentru $i = 1, \ldots, 4$.

Să considerăm evenimentele

$$A = \{\omega_1, \omega_2\},\ B = \{\omega_1, \omega_3\},\ C = \{\omega_1, \omega_4\}.$$

Avem

$$P(A) = P(B) = P(C) = 1/2,$$

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = 1/4 \tag{3.6.3}$$

și

$$P(A \cap B \cap C) = 1/4. \tag{3.6.4}$$

Deoarece

$$P(A)\,P(B) = P(A)\,P(C) = P(B)\,P(C) = 1/4, \tag{3.6.5}$$

rezultă că evenimentele A, B, C sînt independente două cîte două (compară (3.6.3) și (3.6.5)).

Scanned with OKEN Scanner

Deoarece

$$P(A)\,P(B)\,P(C) = 1/8, \tag{3.6.6}$$

rezultă că evenimentele nu sînt independente în totalitate (compară (3.6.4) și (3.6.6)).

Dacă evenimentele $A_1, \ldots, A_n$ sînt independente n cîte n, ele se numesc *independente în totalitate*, sau, mai simplu, *independente*. Evident, dacă $A_1, \ldots, A_n$ sînt independente în totalitate, oricare sistem de m $(m \leqslant n)$ evenimente $A_{i_1}, \ldots, A_{i_m}$ selectate dintre evenimentele $A_1, \ldots, A_n$ este de asemenea un sistem de evenimente independente.

Noțiunea de independență se poate introduce și pentru familii oarecare de evenimente. Mai precis, spunem că familia $(A_i)_{i \in I} \subset \mathcal{K}$ este formată din evenimente independente, dacă pentru orice $J \subset I$, cu J finit, familia $(A_i)_{i \in J}$ este formată din evenimente independente.

Mai mult, noțiunea de independență poate fi extinsă la familii de corpuri boreliene, care, dintr-un punct de vedere general, reprezintă cadrul natural în care trebuie introdusă această noțiune. Mai precis, fie $(E, \mathcal{K}, P)$ un cîmp borelian de probabilitate și fie $(\mathcal{F}_i)_{i \in I}$ o familie de corpuri boreliene pe E cu $\mathcal{F}_i \subset \mathcal{K}$ pentru orice $i \in I$. Spunem că familia $(\mathcal{F}_i)_{i \in I}$ este formată din corpuri boreliene independente dacă, oricare ar fi familia de evenimente $(A_i)_{i \in I}$ cu $A_i \in \mathcal{F}_i$, aceasta este formată din evenimente independente. Din definiție rezultă imediat că dacă $(\mathcal{F}_i)_{i \in I}$ sînt independente și $J \subset I$, atunci $(\mathcal{F}_i)_{i \in J}$ sînt independente. Invers, dacă $(\mathcal{F}_i)_{i \in J}$ sînt independente pentru orice $J \subset I$, cu J finit, rezultă că $(\mathcal{F}_i)_{i \in I}$ sînt independente. În sfîrșit, să mai observăm că, dacă $(\mathcal{F}_i)_{i \in I}$ sînt independente și dacă $\mathcal{F}_i' \subset \mathcal{F}_i$ pentru orice $i \in I$ sînt corpuri boreliene, atunci $(\mathcal{F}_i')_{i \in I}$ sînt de asemenea independente.

Să observăm acum că unei familii de evenimente $(A_i)_{i \in I}$ din $\mathcal{K}$ îi putem asocia familia corpurilor boreliene generate de aceste evenimente, adică $(\mathcal{B}(A_i))_{i \in I}$. Se poate arăta că independența evenimentelor $(A_i)_{i \in I}$ este echivalentă cu independența corpurilor boreliene generate.

Ținînd seama de această observație, putem să introducem și noțiunea de independență a unei familii de desfaceri. Mai

precis, dacă $(\mathfrak{D}_i)_{i \in I}$ este o familie de desfaceri, atunci spunem că aceste desfaceri sînt independente dacă corpurile boreliene generate de aceste desfaceri sînt independente. Se poate arăta că un număr finit de desfaceri $\mathfrak{D}_1, \ldots, \mathfrak{D}_n$ sînt independente, dacă pentru orice evenimente $A_1 \in \mathfrak{D}_1, \ldots, A_n \in \mathfrak{D}_n$, avem

$$P(A_1 \cap \ldots \cap A_n) = P(A_1) \ldots P(A_n).$$

Din cele de mai sus rezultă că putem vorbi de independența unei familii formată din evenimente, desfaceri și corpuri boreliene, înțelegînd prin aceasta independența corpurilor boreliene generate de elementele acestei familii.

Scanned with OKEN Scanner

Capitolul 4

VARIABILE ALEATOARE

4.1. DEFINIȚII ȘI PROPRIETĂȚI

Evenimentelor întîmplătoare li se pot asocia diverse numere. Astfel sînt, de exemplu, numărul de puncte obținut la aruncarea unui zar, numărul de apeluri telefonice la o centrală telefonică într-un interval de timp dat, numărul de băieți din 100 nou-născuți, greutatea unui individ luat la întîmplare din populația unui oraș etc. O *variabilă aleatoare** asociază rezultatelor unui experiment numere reale.

Pentru a defini riguros noțiunea de variabilă aleatoare, să considerăm un cîmp borelian de evenimente $(E, \mathcal{K})$ și cîmpul borelian real $(R, \mathcal{B})$ (vezi § 3.2).

Definiția 4.1.1. *Funcția* $\tilde{x} : E \to R$ *este numită variabilă aleatoare dacă pentru orice* $A \in \mathcal{B}$ *avem* $\tilde{x}^{-1}(A) \in \mathcal{K}$.

Pentru orice $\omega \in E$, vom nota prin x valoarea funcției $\tilde{x}$ în punctul ω, adică $\tilde{x}(\omega) = x$. Convenim să notăm prin aceeași literă atît funcția cît și valoarea ei într-un punct, prezența sau absența semnului $\sim$ deasupra literei respective marcînd distincția necesară.

Să observăm că, dacă $\mathcal{K}$ este corpul borelian total $\mathcal{K}_t = \mathfrak{P}(E)$, atunci orice funcție definită pe E și avînd valori reale este variabilă aleatoare. În cazul cînd $\mathcal{K}$ este corpul borelian banal $\mathcal{K}_0$, atunci singurele variabile aleatoare sînt constantele. Pentru un corp borelian arbitrar $\mathcal{K}$, unele funcții definite pe E și cu valori reale vor fi variabile aleatoare, altele nu vor fi. Deoarece avem totdeauna $\mathcal{K}_0 \subset \mathcal{K}$, rezultă că constantele sînt totdeauna variabile aleatoare.

În limbajul teoriei măsurii, o variabilă aleatoare este o *funcție măsurabilă*. Mai precis, o funcție măsurabilă este definită în modul următor.

---

* Denumirea vine de la cuvîntul latin *alea*, care înseamnă zar, jucător cu zarul, joc de noroc, risc, primejdie.

Scanned with OKEN Scanner

Definiția 4.1.2. *Dacă* $(E, \mathcal{K})$ *și* $(F, \mathcal{L})$ *sînt două spații măsurabile, atunci funcția* $\tilde{x} : E \to F$ *este o funcție* $\mathcal{K}$*-măsurabilă* * *dacă pentru orice* $A \in \mathcal{L}$ *avem* $\tilde{x}^{-1}(A) \in \mathcal{K}$.

Uneori, o funcție măsurabilă mai este numită și *element aleator*.

Teorema 4.1.1. *Fie* $(E, \mathcal{K})$ *și* $(F, \mathcal{L})$ *două spații măsurabile și fie* $\tilde{x}$ *o funcție definită pe* E *și cu valori în* F. *În aceste condiții, avem:*

(a) Familia de evenimente

$$\tilde{x}^{-1}(\mathcal{L}) = \{A : A = \tilde{x}^{-1}(B),\ B \in \mathcal{L}\}$$

este un corp borelian pe E;

(b) Funcția $\tilde{x}$ *este măsurabilă dacă și numai dacă avem*

$$\tilde{x}^{-1}(\mathcal{L}) \subset \mathcal{K}.$$

(c) Pentru orice $\mathcal{M} \subset \mathcal{P}(F)$ *avem*

$$\mathcal{B}(\tilde{x}^{-1}(\mathcal{M})) = \tilde{x}^{-1}(\mathcal{B}(\mathcal{M})).$$

Demonstrație. (a). Dacă $A \in \tilde{x}^{-1}(\mathcal{L})$, atunci $A = \tilde{x}^{-1}(B)$ cu $B \in \mathcal{L}$ și deci

$$A^c = (\tilde{x}^{-1}(B))^c = \tilde{x}^{-1}(B^c) \in \tilde{x}^{-1}(\mathcal{L}),$$

deoarece $B^c \in \mathcal{L}$.

Dacă $(A_i)_{i \in I}$ este o familie cel mult numărabilă din $\mathcal{K}$, atunci există $B_i \in \mathcal{L}$ astfel încît să avem $A_i = \tilde{x}^{-1}(B_i)$ și deci

$$\bigcup_{i \in I} A_i = \bigcup_{i \in I} \tilde{x}^{-1}(B_i) = \tilde{x}^{-1}\left(\bigcup_{i \in I} B_i\right) \in \mathcal{K},$$

deoarece $\bigcup_{i \in I} B_i \in \mathcal{L}$.

În sfîrșit, avem evident $\tilde{x}^{-1}(F) = E \in \mathcal{K}$, deoarece $F \in \mathcal{L}$.

(b). Dacă $\tilde{x}$ este măsurabilă, atunci $\tilde{x}^{-1}(B) \in \mathcal{K}$ pentru orice $B \in \mathcal{L}$, adică $\tilde{x}^{-1}(\mathcal{L}) \subset \mathcal{K}$. Reciproc, dacă $\tilde{x}^{-1}(\mathcal{L}) \subset \mathcal{K}$, atunci, pentru orice $B \in \mathcal{L}$, avem $\tilde{x}^{-1}(B) \in \mathcal{K}$ și deci $\tilde{x}$ este măsurabilă.

---

* Dacă nu există pericol de confuzie se spune, simplu, *măsurabilă*.

Scanned with OKEN Scanner

(c). Din afirmația *(a)* rezultă că $\tilde{x}^{-1}(\mathfrak{B}(\mathfrak{M}))$ este un corp borelian pe E. Dacă $A \in \tilde{x}^{-1}(\mathfrak{M})$, rezultă că există $B \in \mathfrak{M}$ astfel încît să avem $A = \tilde{x}^{-1}(B)$. Dar evident $B \in \mathfrak{B}(\mathfrak{M})$ și deci $A \in \tilde{x}^{-1}(\mathfrak{B}(\mathfrak{M}))$, de unde se deduce $\tilde{x}^{-1}(\mathfrak{M}) \subset \tilde{x}^{-1}(\mathfrak{B}(\mathfrak{M}))$. Deoarece $\mathfrak{B}(\tilde{x}^{-1}(\mathfrak{M}))$ este cel mai mic corp borelian generat de $\tilde{x}^{-1}(\mathfrak{M})$, rezultă $\mathfrak{B}(\tilde{x}^{-1}(\mathfrak{M})) \subset \tilde{x}^{-1}(\mathfrak{B}(\mathfrak{M}))$.

Pentru a demonstra incluziunea inversă, să considerăm familia de evenimente

$$\mathfrak{N} = \{B : \tilde{x}^{-1}(B) \in \mathfrak{B}(\tilde{x}^{-1}(\mathfrak{M}))\}.$$

Evident, dacă $B \in \mathfrak{M}$, atunci $\tilde{x}^{-1}(B) \in \tilde{x}^{-1}(\mathfrak{M})$ și deci $\tilde{x}^{-1}(B) \in B(\tilde{x}^{-1}(\mathfrak{M}))$, ceea ce arată că $B \in \mathfrak{N}$. Cu alte cuvinte, $\mathfrak{M} \subset \mathfrak{N}$.

Pe de altă parte, se constată ușor că $\mathfrak{N}$ este un corp borelian pe F și deci $\mathfrak{B}(\mathfrak{M}) \subset \mathfrak{N}$. Cu alte cuvinte, $B \in \mathfrak{B}(\mathfrak{M})$ implică $\tilde{x}^{-1}(B) \in \mathfrak{B}(\tilde{x}^{-1}(\mathfrak{M}))$, adică $\tilde{x}^{-1}(\mathfrak{B}(\mathfrak{M})) \subset \mathfrak{B}(\tilde{x}^{-1}(\mathfrak{M}))$.

Observație. Dacă $(E, \mathfrak{K})$ și $(F, \mathfrak{L})$ sînt spații măsurabile și $\tilde{x} : E \to F$, atunci din teorema precedentă rezultă că $\tilde{x}^{-1}(\mathfrak{L})$ este cel mai mic corp borelian pentru care $\tilde{x}$ este măsurabilă; corpul borelian $\tilde{x}^{-1}(\mathfrak{L})$ este numit și *corp borelian generat* de $\tilde{x}$ și va fi notat și prin $\mathfrak{B}(\tilde{x})$.

Mai general, dacă $(F_i, \mathfrak{L}_i)_{i \in I}$ este o familie de spații măsurabile și $\tilde{x}_i : E \to F_i$, atunci $\mathfrak{B}\left(\bigcup_{i \in I} \tilde{x}^{-1}(\mathfrak{L}_i)\right)$ este cel mai mic corp borelian pe E pentru care toate funcțiile $\tilde{x}_i$, $i \in I$, sînt măsurabile. El este numit *corpul borelian generat de familia* $(\tilde{x}_i)_{i \in I}$ și este notat și prin $\mathfrak{B}(\tilde{x}_i \mid i \in I)$.

Din teorema precedentă se obține imediat consecința următoare.

Teorema 4.1.2. *Fie* $(\mathrm{E}, \mathfrak{K})$ *un cîmp borelian de evenimente și* $\tilde{\mathrm{x}}$ *o funcție cu valori reale definită pe* E. *Condiția necesară și suficientă ca* $\tilde{\mathrm{x}}$ *să fie o variabilă aleatoare este ca pentru orice* $\mathrm{x} \in \mathrm{R}$ *să avem*

$$\{\omega : \tilde{\mathrm{x}}(\omega) < \mathrm{x}\} \in \mathfrak{K}. \tag{4.1.1}$$

Demonstrație. Necesitatea este imediată, observînd că pentru orice $x \in R$ avem $(-\infty, x) \in \mathfrak{B}$ și deci

$$\tilde{x}^{-1}((-\infty, x)) = \{\omega : \tilde{x}(\omega) < x\} \in \mathfrak{K}.$$

Scanned with OKEN Scanner

Pentru demonstrarea suficienței, să observăm că $\mathfrak{B} = \mathfrak{B}(\mathfrak{M})$, unde

$$\mathfrak{M} = \{(-\infty, x) : x \in R\}.$$

Deoarece, prin ipoteză, $\tilde{x}^{-1}(\mathfrak{M}) \subset \mathfrak{K}$, rezultă $\mathfrak{B}(\tilde{x}^{-1}(\mathfrak{M})) \subset \mathfrak{K}$ și ținînd seama de teorema precedentă $\tilde{x}^{-1}(\mathfrak{B}(\mathfrak{M})) \subset \mathfrak{K}$. Cu alte cuvinte, $\tilde{x}^{-1}(\mathfrak{B}) \subset \mathfrak{K}$, ceea ce arată că $\tilde{x}$ este $\mathfrak{K}$-măsurabilă.

Observație. Condiția (4.1.1) poate fi înlocuită cu una din condițiile

$$\{\omega : \tilde{x}(\omega) \leqslant x\} \in \mathfrak{K};\ \{\omega : \tilde{x}(\omega) > x\} \in \mathfrak{K};\ \{\omega : \tilde{x}(\omega) \geqslant x\} \in \mathfrak{K},$$

teorema precedentă rămînînd adevărată.

În cele ce urmează vom trece în revistă principalele proprietăți ale variabilelor aleatoare.

Teorema 4.1.3. *(a) Dacă* $\tilde{x}$ *și* $\tilde{y}$ *sînt variabile aleatoare pe cîmpul borelian de evenimente* $(E, \mathfrak{K})$, *atunci*

$$\{\omega : \tilde{x}(\omega) < \tilde{y}(\omega)\} \in \mathfrak{K},$$

$$\{\omega : \tilde{x}(\omega) \leqslant \tilde{y}(\omega)\} \in \mathfrak{K},$$

$$\{\omega : \tilde{x}(\omega) = \tilde{y}(\omega)\} \in \mathfrak{K}.$$

(b) Dacă $\tilde{x}$ *și* $\tilde{y}$ *sînt variabile aleatoare pe* $(E, \mathfrak{K})$, *atunci și* $a\tilde{x}$, $|\tilde{x}|$, $\tilde{x}^2$, $\tilde{x}+\tilde{y}$, $\tilde{x}-\tilde{y}$, $\tilde{x}\tilde{y}$ *și* $\tilde{x}/\tilde{y}$ *sînt variabile aleatoare, unde* a *este un număr real arbitrar.*

(c) Dacă $(\tilde{x}_n)$ *este un șir de variabile aleatoare pe* $(E, \mathfrak{K})$, *atunci*

$$\sup_n (\tilde{x}_n);\ \inf_n (\tilde{x}_n);\ \limsup \tilde{x}_n;\ \liminf \tilde{x}_n$$

sînt de asemenea variabile aleatoare pe $(E, \mathfrak{K})$.

(d) Dacă $(E, \mathfrak{K})$, $(F, \mathfrak{L})$ *și* $(G, \mathfrak{M})$ *sînt spații măsurabile și* $\tilde{x} : E \to F$, $\tilde{y} : F \to G$ *sînt funcții măsurabile, atunci* $\tilde{y} \circ \tilde{x}$ *este măsurabilă.*

Demonstrație. *(a)* Dacă notăm prin $(r_n)_{n \in N}$ șirul numerelor raționale, atunci

$$\{\omega : \tilde{x}(\omega) < \tilde{y}(\omega)\} = \bigcup_{n \in N} \{\omega : \tilde{x}(\omega) < r_n\} \cap \{\omega : \tilde{y}(\omega) > r_n\},$$

de unde rezultă prima apartenență. Ținînd seama că

$$\{\omega : \tilde{x}(\omega) \leqslant \tilde{y}(\omega)\} = \{\omega : \tilde{x}(\omega) > \tilde{y}(\omega)\}^c$$

și

$$\{\omega : \tilde{x}(\omega) = \tilde{y}(\omega)\} = \{\omega : \tilde{x}(\omega) \leqslant \tilde{y}(\omega)\} \cap \{\omega : \tilde{x}(\omega) \geqslant \tilde{y}(\omega)\},$$

rezultă și celelalte două apartenențe.

(b) Să arătăm mai întîi că $a\tilde{x}$ este variabilă aleatoare. Dacă $a = 0$, rezultatul este evident. Dacă $a > 0$, atunci pentru orice $x \in R$ avem

$$\{\omega : a\tilde{x}(\omega) < x\} = \{\omega : \tilde{x}(\omega) < x/a\} \in \mathcal{K}$$

și deci $a\tilde{x}$ este variabilă aleatoare. Dacă $a < 0$, rezultatul se obține în mod analog.

Pentru a arăta că $|\tilde{x}|$ este variabilă aleatoare, să observăm că pentru orice $x < 0$ avem

$$\{\omega : |\tilde{x}(\omega)| > x\} = E \in \mathcal{K}$$

și pentru orice $x \geqslant 0$,

$$\{\omega : |\tilde{x}(\omega)| > x\} = \{\omega : \tilde{x}(\omega) > x\} \cup \{\omega : \tilde{x}(\omega) < -x\} \in \mathcal{K},$$

de unde rezultă că $|\tilde{x}|$ este variabilă aleatoare.

Pentru a arăta că $\tilde{x}^2$ este variabilă aleatoare, să observăm că pentru orice $x < 0$ avem

$$\{\omega : \tilde{x}^2(\omega) > x\} = E \in \mathcal{K},$$

iar pentru orice $x \geqslant 0$,

$$\{\omega : \tilde{x}^2(\omega) > x\} = \{\omega : |\tilde{x}(\omega)| > \sqrt{x}\} \in \mathcal{K},$$

de unde reiese că $\tilde{x}^2$ este măsurabilă.

Pentru a arăta că $\tilde{x} - \tilde{y}$ este măsurabilă, să observăm că pentru orice $x \in R$ avem

$$\{\omega : \tilde{x}(\omega) - \tilde{y}(\omega) < x\} = \{\omega : \tilde{x}(\omega) < \tilde{y}(\omega) + x\}$$

și deci, ținînd seama de rezultatul de la *(a)*, obținem

$$\{\omega : \tilde{x}(\omega) - \tilde{y}(\omega) < x\} \in \mathcal{K},$$

ceea ce arată că $\tilde{x} - \tilde{y}$ este variabilă aleatoare.

Scanned with OKEN Scanner

Întrucît $\tilde{x}+\tilde{y}=\tilde{x}-(-\tilde{y})$, rezultă că și $\tilde{x}+\tilde{y}$ este variabilă aleatoare.

Deoarece

$$\tilde{x}\tilde{y}=((\tilde{x}+\tilde{y})^2-(\tilde{x}-\tilde{y})^2)/4,$$

se deduce ușor că $\tilde{x}\tilde{y}$ este de asemenea variabilă aleatoare.

Să presupunem că $\tilde{y}(\omega)\neq 0$ pentru orice $\omega\in E$ și să arătăm că $1/\tilde{y}$ este variabilă aleatoare. Într-adevăr, pentru orice $x>0$ avem

$$\{\omega : 1/\tilde{y}(\omega)>x\}=\{\omega : \tilde{y}(\omega)>0\}\cap\{\omega : \tilde{y}(\omega)<1/x\}\in\mathfrak{K}.$$

Pentru orice $x<0$ avem

$$\{\omega : 1/\tilde{y}(\omega)>x\}=$$

$$=\{\omega : \tilde{y}(\omega)>0\}\cup\{\omega : \tilde{y}(\omega)<x\}\cap\{\omega : \tilde{y}(\omega)<1/x\}\in\mathfrak{K}.$$

În sfîrșit, pentru $x=0$, obținem

$$\{\omega : 1/\tilde{y}(\omega)>x\}=\{\omega : \tilde{y}(\omega)>0\}\in\mathfrak{K}.$$

Din cele de mai sus rezultă că $1/\tilde{y}$ este variabilă aleatoare. Deoarece $\tilde{x}/\tilde{y}=\tilde{x}(1/\tilde{y})$, rezultă că $\tilde{x}/\tilde{y}$ este variabilă aleatoare.

(c) Pentru orice $x\in R$ avem

$$\{\omega : \sup_n(\tilde{x}_n(\omega))>x\}=\bigcup_{n\in N}\{\omega : \tilde{x}_n(\omega)>x\}\in\mathfrak{K}$$

și deci $\sup(\tilde{x}_n)$ este variabilă aleatoare. Deoarece

$$\inf_n(\tilde{x}_n)=-\sup_n(-\tilde{x}_n),$$

rezultă că și $\inf(\tilde{x}_n)$ este variabilă aleatoare.

Întrucît

$$\lim_n\sup\tilde{x}_n=\inf_{n\geq 1}\sup_{m\geq n}(\tilde{x}_m)$$

și

$$\lim_n\inf\tilde{x}_n=\sup_{n\geq 1}\inf_{m\geq n}(\tilde{x}_m),$$

se deduc imediat și celelalte două afirmații.

(d) Deoarece

$$(\tilde{y}\circ\tilde{x})^{-1} = \tilde{x}^{-1}\circ\tilde{y}^{-1},$$

rezultă că pentru orice $A \in \mathfrak{M}$ avem

$$(\tilde{y}\circ\tilde{x})^{-1}(A) = \tilde{x}^{-1}(\tilde{y}^{-1}(A)) \in \mathfrak{K},$$

ținînd seama că $\tilde{y}^{-1}(A) \in \mathfrak{L}$.

Definiția 4.1.3. Vom spune că o variabilă aleatoare este *simplă* dacă poate lua numai un număr finit de valori.

Cel mai simplu exemplu de variabilă aleatoare simplă este funcția constantă pe E. Un alt exemplu de variabilă aleatoare simplă este *indicatorul* χ_A al unui eveniment $A \in \mathfrak{K}$, definit astfel

$$\chi_A(\omega) = \begin{cases} 1 & \text{dacă} \quad \omega \in A, \\ 0 & \text{dacă} \quad \omega \notin A. \end{cases} \tag{4.1.2}$$

Dacă variabila aleatoare simplă $\tilde{x}$ ia valorile $x_1, \ldots, x_n$ și notăm $A_j = \tilde{x}^{-1}(\{x_j\}) \in \mathfrak{K}$ pentru $1 \leqslant j \leqslant n$, atunci avem evident $A_i \cap A_j = \varnothing$ pentru $i \neq j$ și $A_1 \cup \ldots \cup A_n = E$; cu alte cuvinte, evenimentele $A_1, \ldots, A_n$ realizează o desfacere a evenimentului sigur E. Este atunci clar că variabila aleatoare simplă $\tilde{x}$ poate fi scrisă sub forma

$$\tilde{x}(\omega) = x_1\chi_{A_1}(\omega) + \ldots + x_n\chi_{A_n}(\omega). \tag{4.1.3}$$

Se poate vedea fără dificultate că dacă $\tilde{x}$ și $\tilde{y}$ sînt variabile aleatoare simple, atunci $|\tilde{x}|$, $\tilde{x}+\tilde{y}$, $\tilde{x}-\tilde{y}$, $\tilde{x}\tilde{y}$ sînt de asemenea variabile aleatoare simple.

Teorema 4.1.4. *Fie* $\tilde{\mathrm{x}}$ *o variabilă aleatoare simplă pe* (E, $\mathfrak{K}$). *Există un șir crescător de variabile aleatoare simple* $(\tilde{\mathrm{x}}_n)_{n\in\mathrm{N}}$ *pe* (E, $\mathfrak{K}$) *astfel încît să avem*

$$\lim_{\mathrm{n}} \tilde{\mathrm{x}}_n(\omega) = \tilde{\mathrm{x}}(\omega) \tag{4.1.4}$$

pentru orice $\omega \in \mathrm{E}$.

Demonstrație. Să considerăm mai întîi cazul în care $\tilde{x}(\omega) \geqslant 0$ pentru orice $\omega \in E$. Construim șirul $(\tilde{x}_n)_{n\in N}$ în modul următor:

$$\tilde{x}_n(\omega) = \begin{cases} (i-1)/2^n & \text{dacă } (i-1)/2^n \leqslant \tilde{x}(\omega) < i/2^n, \ 1 \leqslant i \leqslant n2^n, \\ n & \text{dacă } \tilde{x}(\omega) \geqslant n. \end{cases}$$

Scanned with OKEN Scanner

Este clar că fiecare variabilă aleatoare $\tilde{x}_n$ este o variabilă aleatoare simplă, iar șirul $(\tilde{x}_n(\omega))_{n\in N}$ este crescător pentru fiecare $\omega\in E$.

Dacă pentru un anumit $\omega\in E$ avem $\tilde{x}(\omega) < +\infty$, atunci

$$0 \leqslant \tilde{x}(\omega) - \tilde{x}_n(\omega) \leqslant i/2^n - (i-1)/2^n = 1/2^n$$

și deci se obține relația (4.1.4).

Dacă pentru un anumit $\omega\in E$ avem $\tilde{x}(\omega) = +\infty$, rezultă că $\tilde{x}_n(\omega) = n$ pentru orice $n\in N$ și deci $\tilde{x}_n(\omega) \to +\infty$, obținîndu-se din nou afirmația (4.1.4).

Dacă variabila aleatoare $\tilde{x}$ nu este nenegativă, vom scrie

$$\tilde{x}(\omega) = \tilde{x}^+(\omega) - \tilde{x}^-(\omega),$$

unde

$$\tilde{x}^+(\omega) = \max(\tilde{x}(\omega), 0);\ \tilde{x}^-(\omega) = \min(\tilde{x}(\omega), 0)$$

sînt evident variabile aleatoare nenegative. Dacă $(\tilde{x}_n^+)$ și $(\tilde{x}_n^-)$ sînt șirurile de variabile aleatoare simple care converg respectiv către $\tilde{x}^+$ și $\tilde{x}^-$, atunci șirul $(\tilde{x}_n)$ definit de

$$\tilde{x}_n = \tilde{x}_n^+ - \tilde{x}_n^-$$

este evident un șir de variabile aleatoare simple convergent spre $\tilde{x}$.

4.2. FUNCȚIA DE REPARTIȚIE A UNEI VARIABILE ALEATOARE

Să considerăm un cîmp borelian de probabilitate $(E, \mathcal{K}, P)$, o mulțime arbitrară F și o funcție $\tilde{x} : E \to F$. Ne punem problema organizării lui F ca un cîmp borelian de probabilitate $(F, \mathcal{K}', P')$, utilizînd în acest scop aplicația $\tilde{x}$.

Pentru aceasta vom lua drept $\mathcal{K}'$ mulțimea

$$\mathcal{K}' = \{A : A\in\mathcal{P}(\mathcal{F}),\ \tilde{x}^{-1}(A)\in\mathcal{K}\}.$$

Este ușor de verificat că familia $\mathcal{K}'$ astfel definită este un corp borelian pe F. Într-adevăr, dacă $A\in\mathcal{K}'$, rezultă $\tilde{x}^{-1}(A)\in\mathcal{K}$ și deci $(\tilde{x}^{-1}(A))^c\in\mathcal{K}$ sau $\tilde{x}^{-1}(A^c)\in\mathcal{K}$, adică $A^c\in\mathcal{K}'$. În conti-

nuare, dacă $(A_i)_{i\in I}$ este o familie cel mult numărabilă de elemente din $\mathcal{K}'$, rezultă că $\tilde{x}^{-1}(A_i)\in\mathcal{K}$ pentru orice $i\in I$ și deci $\bigcup_{i\in I}\tilde{x}^{-1}(A_i)\in\mathcal{K}$ sau $\tilde{x}^{-1}\left(\bigcup_{i\in I}A_i\right)\in\mathcal{K}$, de unde obținem $\bigcup_{i\in I}A_i\in\mathcal{K}'$. În sfîrșit, avem $\tilde{x}^{-1}(F)=E\in\mathcal{K}$ și deci $F\in\mathcal{K}'$. Corpul borelian $\mathcal{K}'$ va fi notat prin $\mathcal{K}_{\tilde{x}}$.

În al doilea rînd, vom lua drept P' funcția cu valori reale definită pe $\mathcal{K}_{\tilde{x}}$ în modul următor: pentru orice $A\in\mathcal{K}_{\tilde{x}}$ punem $P'(A)=P(\tilde{x}^{-1}(A))$. Se poate verifica ușor că P' este o probabilitate complet aditivă pe $\mathcal{K}_{\tilde{x}}$. Probabilitatea P' astfel introdusă este numită și *funcția de probabilitate a lui* $\tilde{x}$ sau *repartiția* lui $\tilde{x}$ și va fi notată prin $P_{\tilde{x}}$ sau prin $P\circ\tilde{x}^{-1}$.

Teorema 4.2.1. *Fie* $(E,\mathcal{K})$ *și* $(F,\mathfrak{L})$ *două spații măsurabile. Funcția* $\tilde{x}:E\to F$ *este măsurabilă dacă și numai dacă avem* $\mathcal{K}_{\tilde{x}}\supset\mathfrak{L}$.

Demonstrație. Dacă $\tilde{x}$ este măsurabilă, atunci pentru orice $A\in\mathfrak{L}$ avem $\tilde{x}^{-1}(A)\in\mathcal{K}$ și deci $A\in\mathcal{K}_{\tilde{x}}$, adică $\mathfrak{L}\subset\mathcal{K}_{\tilde{x}}$.

Reciproc, dacă $\mathfrak{L}\subset\mathcal{K}_{\tilde{x}}$, rezultă că pentru orice $A\in\mathfrak{L}$ avem $A\in\mathcal{K}_{\tilde{x}}$, adică $\tilde{x}^{-1}(A)\in\mathcal{K}$ și deci $\tilde{x}$ este măsurabilă.

Definiția 4.2.1. *Fie* $\tilde{x}$ *o variabilă aleatoare pe cîmpul borelian de probabilitate* $(E,\mathcal{K},P)$. *Funcția de variabilă reală* F *definită de*

$$F(x)=P_{\tilde{x}}((-\infty,x))$$

pentru orice $x\in R$ *este numită funcția de repartiție a lui* $\tilde{x}$.

Teorema 4.2.2. *Dacă* $\tilde{x}$ *este o variabilă aleatoare pe cîmpul borelian de probabilitate* $(E,\mathcal{K},P)$ *și dacă* F *este funcția sa de repartiție, atunci:*

(a) Avem $F(-\infty)=0$ *și* $F(+\infty)=1$;

(b) Funcția F *este crescătoare pe* R;

(c) Funcția F *este continuă la stînga pe* R.

Demonstrație. (a). Să considerăm șirul (x_n) de numere reale descrescător cu limita $-\infty$. Șirul de evenimente $(A_n)\subset\mathfrak{B}$ definit de $A_n=(-\infty,x_n)$ este descrescător și $\bigcap_{n\in N}A_n=\varnothing$. Din proprietatea de continuitate a probabilității $P_{\tilde{x}}$ obținem

$$\lim_n P_{\tilde{x}}(A_n)=P_{\tilde{x}}\left(\bigcap_{n\in N}A_n\right),$$

adică

$$\lim_n F(x_n) = P_{\tilde{x}}(\varnothing) = 0,$$

de unde $F(-\infty) = 0$.

Pentru a demonstra a doua parte a afirmației, fie (y_n) un șir de numere reale crescător cu limita $+\infty$. Șirul de evenimente $(B_n) \subset \mathfrak{B}$ definit de $B_n = (-\infty, y_n)$ este crescător și $\bigcup_{n \in N} B_n = R$. Aplicînd din nou proprietatea de continuitate a probabilității $P_{\tilde{x}}$ obținem

$$\lim_n P_{\tilde{x}}(B_n) = P_{\tilde{x}}\left(\bigcup_{n \in N} B_n\right),$$

adică

$$\lim_n F(y_n) = P_{\tilde{x}}(R) = 1,$$

de unde $F(+\infty) = 1$.

(b). Pentru orice $x, y \in R$, cu $x < y$, avem $(-\infty, x) \subset (-\infty, y)$ și deci

$$P_{\tilde{x}}((-\infty, x)) \leqslant P_{\tilde{x}}((-\infty, y)),$$

adică $F(x) \leqslant F(y)$, ceea ce arată că F este crescătoare pe R.

(c). Fie $x \in R$ și fie (x_n) un șir de numere reale crescător cu limita x. Fie $(C_n) \subset \mathfrak{B}$ șirul de evenimente definit de $C_n = [x_n, x)$. Este clar că șirul (C_n) este descrescător și că $\bigcap_{n \in N} C_n = \varnothing$. Aplicînd din nou proprietatea de continuitate a probabilității $P_{\tilde{x}}$ rezultă

$$\lim_n P_{\tilde{x}}(C_n) = P_{\tilde{x}}(\varnothing) = 0.$$

Deoarece

$$P_{\tilde{x}}(C_n) = P_{\tilde{x}}((-\infty, x) - (-\infty, x_n)) =$$

$$= P_{\tilde{x}}((-\infty, x)) - P_{\tilde{x}}((-\infty, x_n)) = F(x) - F(x_n),$$

rezultă că

$$\lim_n F(x_n) = F(x)$$

adică $F(x-0) = F(x)$, ceea ce arată că F este continuă la stînga în $x \in R$. Dar cum x a fost arbitrar, rezultă că F este continuă la stînga pe R.

Scanned with OKEN Scanner

Observația 1. Se poate arăta că proprietățile *(a)*, *(b)* și *(c)* din teorema precedentă caracterizează funcțiile de repartiție. Cu alte cuvinte, orice funcție care are proprietățile *(a)*, *(b)* și *(c)* este funcția de repartiție a unei anumite variabile aleatoare. Trebuie observat însă că mulțimea variabilelor aleatoare și mulțimea funcțiilor de repartiție nu sînt în corespondență bijectivă. Mai precis, unei variabile aleatoare îi corespunde o singură funcție de repartiție, pe cînd o funcție de repartiție poate fi asociată cu mai multe variabile aleatoare.

Observația 2. Din definiția funcției de repartiție rezultă că $F(x)$ reprezintă probabilitatea evenimentului $\{\omega : \tilde{x}(\omega) < x\}$ sau, mai pe scurt, probabilitatea ca $\tilde{x} < x$. În practică ne interesează și exprimarea probabilității evenimentelor $\{\omega : \tilde{x}(\omega) \leqslant x\}$ și $\{\omega : \tilde{x}(\omega) = x\}$ cu ajutorul funcției de repartiție F. Pentru a calcula prima probabilitate, să considerăm șirul de evenimente $(D_n) \subset \mathfrak{B}$ definit de $D_n = (-\infty, x+1/n)$. Este clar că șirul (D_n) este descrescător și că $\bigcap_{n \in N} D_n = (-\infty, x]$. Aplicînd încă o dată proprietatea de continuitate a probabilității $P_{\tilde{x}}$ obținem

$$\lim_n P_{\tilde{x}}(D_n) = P_{\tilde{x}}\left(\bigcap_{n \in N} D_n\right)$$

sau

$$\lim_n F(x+1/n) = P_{\tilde{x}}((-\infty, x]) = P(\{\omega : \tilde{x}(\omega) \leqslant x\})$$

și deci

$$P(\{\omega : \tilde{x}(\omega) \leqslant x\}) = F(x+0). \tag{4.2.1}$$

Pentru calculul probabilității celui de al doilea eveniment, să observăm că

$$\{\omega : \tilde{x}(\omega) = x\} = \{\omega : \tilde{x}(\omega) \leqslant x\} - \{\omega : \tilde{x}(\omega) < x\}$$

și deci

$$P(\tilde{x} = x) = P(\tilde{x} \leqslant x) - P(\tilde{x} < x),$$

adică

$$P(\tilde{x} = x) = F(x+0) - F(x-0), \tag{4.2.2}$$

ținînd seama și de faptul că $F(x) = F(x-0)$.

În sfîrșit, dacă $x, y \in R$ cu $x < y$, atunci din

$$[x, y) = (-\infty, y) - (-\infty, x)$$

rezultă

$$P_{\tilde{x}}([x, y)) = P_{\tilde{x}}((-\infty, y)) - P_{\tilde{x}}((-\infty, x)),$$

adică

$$P(x \leqslant \tilde{x} < y) = F(y) - F(x). \tag{4.2.3}$$

Observația 3. Uneori în aplicații este util să cunoaștem *funcția de repartiție reciprocă* G a variabilei aleatoare $\tilde{x}$, definită prin

$$G(x) = 1 - F(x) = P(\{\omega : \tilde{x}(\omega) \geqslant x\}),$$

pentru orice $x \in R$. Este clar că G are proprietățile următoare:

(a') $G(-\infty) = 1$ *și* $G(+\infty) = 0$;

(b') *Funcția* G *este descrescătoare pe* R;

(c') *Funcția* G *este continuă la stînga pe* R.

Utilizînd funcția de repartiție reciprocă, formulele (4.2.1), (4.2.2) și (4.2.3) devin:

$$P(\tilde{x} \leqslant x) = 1 - G(x+0); \tag{4.2.4}$$

$$P(\tilde{x} = x) = G(x-0) - G(x+0); \tag{4.2.5}$$

$$P(x \leqslant \tilde{x} < y) = G(x) - G(y). \tag{4.2.6}$$

Vom prezenta în continuare alte cîteva proprietăți importante ale funcției de repartiție.

Teorema 4.2.3. *Mulțimea punctelor de discontinuitate ale unei funcții de repartiție este cel mult numărabilă.*

Demonstrație. Fie F o funcție de repartiție. *Saltul* funcției F în punctul $x \in R$ este prin definiție $F(x+0) - F(x-0)$. Deoarece $0 \leqslant F(x) \leqslant 1$, funcția F are cel mult un salt mai mare decît 1/2, cel mult trei salturi cuprinse între $1/2^2$ și 1/2 și, în general, cel mult $2^n - 1$ salturi cuprinse între $1/2^n$ și $1/2^{n-1}$. Mulțimea salturilor se poate deci numerota, prin urmare ea este o mulțime cel mult numărabilă.

Scanned with OKEN Scanner

Teorema 4.2.4. *O funcție de repartiție* F *continuă pe* R *este uniform continuă pe* R.

Demonstrație. Deoarece $F(-\infty) = 0 \leqslant F(x) \leqslant F(+\infty) = 1$, pentru orice $x \in R$, rezultă că pentru orice $\varepsilon > 0$ dat, putem determina numerele reale $a(\varepsilon)$ și $b(\varepsilon)$ cu proprietățile următoare:

$$0 \leqslant F(x) \leqslant \varepsilon/2 \text{ pentru } x \in (-\infty, a(\varepsilon))$$

și

$$1 - \varepsilon/2 \leqslant F(x) \leqslant 1 \text{ pentru } x \in (b(\varepsilon), +\infty).$$

Din continuitatea lui F pe intervalul închis $[a(\varepsilon), b(\varepsilon)]$ rezultă că ea este uniform continuă pe acest interval. Se poate deci determina un număr $\eta(\varepsilon)$ cu proprietatea că

$$|F(x) - F(y)| \leqslant \varepsilon/2$$

pentru orice $x, y \in [a(\varepsilon), b(\varepsilon)]$ pentru care $|x - y| < \eta(\varepsilon)$.

Dacă notăm prin $\delta(\varepsilon) = \min(\eta(\varepsilon), b(\varepsilon) - a(\varepsilon))$, din cele de mai sus rezultă

$$|F(x) - F(y)| < \varepsilon$$

pentru orice $x, y \in R$ pentru care $|x - y| < \delta(\varepsilon)$. Așadar, F este uniform continuă pe R.

În continuare vom arăta că orice funcție de repartiție se descompune într-o componentă continuă și într-o componentă discontinuă.

Conform teoremei 4.2.3, orice funcție de repartiție F are o mulțime cel mult numărabilă de discontinuități (salturi). Fie $\{x_n\}$ această mulțime. Saltul p_n al funcției de repartiție în punctul x_n este prin definiție

$$p_n = F(x_n + 0) - F(x_n - 0).$$

Să considerăm funcția în scară $S(x)$ definită astfel

$$S(x) = \sum_{\{n\,:\,x_n < x\}} p_n.$$

Se poate vedea ușor că $S(x)$ este o funcție crescătoare și continuă la stînga cu proprietatea că $S(-\infty) = 0$. Vom nota $S(+\infty) = a$, unde $0 < a \leqslant 1$.

Scanned with OKEN Scanner

Să considerăm acum funcția C de variabilă reală definită pentru orice $x \in R$ prin

$$C(x) = F(x) - S(x).$$

Este clar că $C(-\infty) = 0$ și că $C(x)$ este o funcție continuă la stînga; vom nota $C(+\infty) = 1 - a = b$.

Se poate arăta că funcția $C(x)$ este continuă la dreapta și deci este o funcție continuă pe R. Într-adevăr, pentru orice $x < y$ avem

$$C(y) - C(x) = F(y) - F(x) - \sum_{x \leqslant x_n < y} p_n.$$

Dacă x este punct de continuitate pentru F, atunci $F(x) = F(x+0)$ și deci

$$C(y) - C(x) = F(y) - F(x+0) - \sum_{x < x_n < y} p_n.$$

Dacă x este punct de discontinuitate pentru F, atunci putem scrie

$$F(y) - F(x) - \sum_{x \leqslant x_n < y} p_n =$$

$$= F(y) - F(x) - F(x+0) + F(x-0) - \sum_{x < x_n < y} p_n,$$

și deci obținem din nou

$$C(y) - C(x) = F(y) - F(x+0) - \sum_{x < x_n < y} p_n.$$

Dacă trecem la limită pentru $y \to x$, $y > x$, rezultă $C(x+0) = C(x)$ și prin urmare C este o funcție continuă pe R.

Rezultă așadar că funcția de repartiție F se scrie ca suma dintre o funcție în scară $S(x)$ și o funcție continuă $C(x)$.

Să arătăm că această descompunere este unică. Într-adevăr, dacă ar mai exista o altă descompunere

$$F(x) = S_1(x) + C_1(x),$$

atunci am obține

$$S_1(x) - S(x) = C_1(x) - C(x),$$

de unde rezultă că ambii membri ai egalității sînt identic nuli (în stînga avem o funcție în scară, iar în dreapta o funcție continuă).

Scanned with OKEN Scanner

Dacă facem acum notația

$$F_d(x) = \frac{1}{a} S(x); \quad F_c(x) = \frac{1}{b} C(x),$$

obținem următorul rezultat.

Teorema 4.2.5. *Orice funcție de repartiție* F *poate fi descompusă în modul următor:*

$$F(x) = aF_d(x) + bF_c(x), \tag{4.2.7}$$

unde F_d *este o funcție de repartiție în scară,* F_c *este o funcție de repartiție continuă, iar* a *și* b *sînt cuprinși între zero și unu și* $a+b = 1$.

Se poate merge cu această descompunere mai departe. Mai exact, din teorema de descompunere a lui Lebesgue rezultă că funcția de repartiție F_c poate fi scrisă sub forma

$$F_c(x) = cF_{ac}(x) + dF_s(x), \tag{4.2.8}$$

unde c și d sînt nenegativi și $c+d = 1$, F_{ac} este o funcție de repartiție absolut continuă, iar F_s este o funcție de repartiție continuă singulară (adică o funcție continuă ale cărei puncte de creștere aparțin unei mulțimi de măsură Lebesgue nulă). Mai exact, F_{ac} este o funcție de repartiție care se poate scrie sub forma

$$F_{ac}(x) = \int_{-\infty}^{x} f(t)\, dt, \tag{4.2.9}$$

unde f este o funcție măsurabilă nenegativă.

Din relațiile (4.2.7) și (4.2.8) rezultă în definitiv că avem

$$F(x) = aF_d(x) + b(cF_{ac}(x) + dF_s(x)).$$

Dacă facem notația $a_1 = a$, $a_2 = bc$, $a_3 = bd$, rezultă

$$F(x) = a_1F_d(x) + a_2F_{ac}(x) + a_3F_s(x), \tag{4.2.10}$$

unde a_1, a_2 și a_3 sînt numere nenegative și $a_1+a_2+a_3 = 1$.

Din cele de mai sus rezultă teorema următoare.

Scanned with OKEN Scanner

Teorema 4.2.6. *Orice funcție de repartiție* F *este o combinație convexă a trei funcții de repartiție: o funcție de repartiție în scară* F_d, *o funcție de repartiție absolut continuă* F_{ac} *și o funcție de repartiție continuă singulară* F_s.

În aplicațiile practice se utilizează în mod obișnuit două clase de variabile aleatoare: *variabile aleatoare discrete* și *variabile aleatoare continue*.

O variabilă aleatoare $\tilde{x}$ este numită *discretă* dacă ia o mulțime cel mult numărabilă de valori. Dacă $(x_i)_{i\in I}$ este mulțimea valorilor lui $\tilde{x}$, atunci probabilitățile

$$p_i = P(\tilde{x} = x_i) = F(x_i + 0) - F(x_i - 0), \quad i \in I, \tag{4.2.11}$$

determină complet funcția de repartiție a lui $\tilde{x}$, care este dată de

$$F(x) = \sum_{\{i\in I : x_i < x\}} p_i. \tag{4.2.12}$$

Cu alte cuvinte, funcția de repartiție a unei variabile aleatoare discrete coincide cu componenta sa în scară (în formula (4.2.10) avem deci $a_1 = 1$, $a_2 = a_3 = 0$); o funcție de repartiție în scară va fi numită și *funcție de repartiție discretă*.

Deoarece o variabilă discretă $\tilde{x}$ cu mulțimea valorilor $(x_i)_{i\in I}$ este complet descrisă de probabilitățile $(p_i)_{i\in I}$ definite de (4.2.11), rezultă că putem utiliza direct aceste probabilități în locul funcției de repartiție dată de (4.2.12). Sîntem astfel conduși la introducerea noțiunii de *funcție de masă* (sau *funcție de frecvență*) asociată unei variabile aleatoare discrete.

Numim *funcție de masă* a variabilei aleatoare discrete $\tilde{x}$ cu mulțimea valorilor $(x_i)_{i\in I}$ funcția f de variabilă reală definită prin

$$f(x) = \begin{cases} 0 & \text{pentru } x \notin (x_i)_{i\in I} \\ p_i & \text{pentru } x = x_i, \ i \in I, \end{cases} \tag{4.2.13}$$

unde p_i, $i \in I$, sînt date de relația (4.2.11). Funcția de repartiție F corespunzătoare se poate exprima cu ajutorul funcției de masă f în modul următor:

$$F(x) = \sum_{\{y : y < x\}} f(y). \tag{4.2.14}$$

Scanned with OKEN Scanner

Este clar că funcția de masă f are proprietățile următoare:

(a) $f(x) \geqslant 0$ *pentru orice* $x \in R$;

(b) $f(x) = 0$ *pentru orice* $x \notin (x_i)_{i \in I}$;

(c) $\sum_{i \in I} f(x_i) = 1$.

Se poate vedea ușor că aceste proprietăți sînt caracteristice pentru funcția de masă, adică orice funcție f cu proprietățile *(a)*, *(b)* și *(c)* de mai sus este funcția de masă a unei anumite variabile aleatoare discrete.

Ținînd seama de formula (4.2.14), rezultă că

$$P(\tilde{x} = x) = f(x),$$

$$P(x \leqslant \tilde{x} < y) = \sum_{\{t: x \leqslant t < y\}} f(t),$$

$$P(\tilde{x} \leqslant x) = \sum_{\{t: t \leqslant x\}} f(t).$$

Analog se pot exprima probabilitățile ca $\tilde{x}$ să aparțină altor tipuri de intervale.

Unei variabile aleatoare discrete $\tilde{x}$ cu mulțimea valorilor $(x_i)_{i \in I}$ îi putem asocia un arbore, ale cărui muchii corespund valorilor x_i, $i \in I$, și probabilităților p_i, $i \in I$ (vezi fig. 4.2.1).

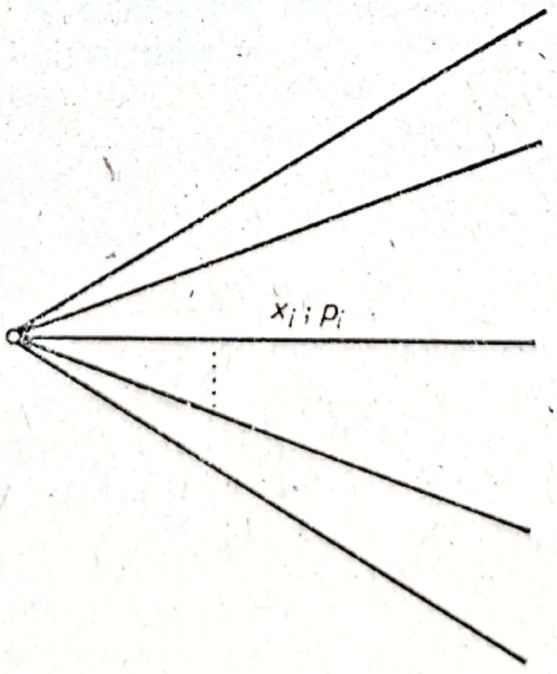


Fig. 4.2.1

Scanned with OKEN Scanner

Aceeași variabilă aleatoare poate fi reprezentată și prin tabloul

$$\tilde{x} : \begin{pmatrix} x_i \\ p_i \end{pmatrix} i \in I,$$

care este numit și *repartiția variabilei discrete* $\tilde{x}$ și care conține valorile și probabilitățile cu care aceste valori sînt luate de către variabila aleatoare.

Să trecem acum la cealaltă clasă importantă de variabile aleatoare. Spunem că o variabilă aleatoare $\tilde{x}$ este *continuă* dacă funcția sa de repartiție F este absolut continuă. Rezultă deci că funcția de repartiție a unei variabile aleatoare continue coincide cu componenta sa absolut continuă (în formula (4.2.10) avem $a_1 = a_3 = 0$ și $a_2 = 1$).

Prima observație pe care o facem este că probabilitatea ca variabila aleatoare continuă $\tilde{x}$ să ia o valoare $x \in R$ este nulă. Într-adevăr,

$$P(\tilde{x} = x) = F(x+0) - F(x-0) = 0,$$

cum rezultă din continuitatea lui F. Aceasta înseamnă că noțiunea de funcție de masă introdusă mai sus pentru variabilele aleatoare discrete nu este adecvată pentru variabilele aleatoare continue. În acest caz vom introduce o altă caracterizare, bazată pe noțiunea de *densitate de probabilitate*.

După cum știm, funcția de repartiție F a unei variabile aleatoare continue $\tilde{x}$ este absolut continuă, adică admite reprezentarea

$$F(x) = \int_{-\infty}^{x} f(t)\, dt, \tag{4.2.15}$$

unde f este o funcție măsurabilă și nenegativă. Funcția f care apare în acest mod este numită *densitatea de probabilitate* a variabilei aleatoare $\tilde{x}$. Din proprietățile funcției de repartiție rezultă ușor că densitatea de probabilitate are proprietățile:

(DP_1) $f(x) \geqslant 0$ *pentru orice* $x \in R$;

$$(DP_2) \int_{-\infty}^{+\infty} f(x)\, dx = 1.$$

Scanned with OKEN Scanner

Se poate arăta că proprietățile (DP_1) și (DP_2) ale densității de probabilitate sînt caracteristice. Cu alte cuvinte, orice funcție măsurabilă f care are aceste proprietăți este densitatea de probabilitate a unei anumite variabile aleatoare continue.

Din reprezentarea (4.2.15) rezultă imediat că

$$\frac{\mathrm{d}F(x)}{\mathrm{d}x} = f(x) \tag{4.2.16}$$

aproape peste tot în raport cu măsura Lebesgue.

O proprietate interesantă a densității de probabilitate se obține din relația

$$f(x) = \lim_{h \to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \to 0} \frac{P(x \leqslant \tilde{x} \leqslant x+h)}{h}$$

sau

$$P(x \leqslant \tilde{x} \leqslant x + \mathrm{d}x) \simeq f(x)\,\mathrm{d}x. \tag{4.2.17}$$

Ultima relație ne arată că produsul $f(x)\mathrm{d}x$ este aproximativ egal cu probabilitatea ca $\tilde{x}$ să ia valori în intervalul $[x, x+\mathrm{d}x]$ de lungime mică; produsul $f(x)\mathrm{d}x$ este numit uneori și *probabilitatea elementară*.

În încheiere să observăm că pentru o variabilă aleatoare discretă $\tilde{x}$ cu mulțimea valorilor $(x_i)_{i \in I}$, funcția de masă $f(x)$ ia valori nenule pe o mulțime cel mult numărabilă. O valoare $x \in R$ va fi numită *valoare posibilă* a lui $\tilde{x}$ dacă avem $f(x) > 0$. Evident, mulțimea valorilor posibile ale lui $\tilde{x}$ este inclusă în mulțimea tuturor valorilor $(x_i)_{i \in I}$.

Dacă $\tilde{x}$ este o variabilă aleatoare continuă, atunci am notat prin $f(x)$ densitatea de probabilitate a lui $\tilde{x}$. Vom spune că $x \in R$ este o *valoare posibilă* a lui $\tilde{x}$ dacă $f(x) > 0$. Evident, $\tilde{x}$ ia valoarea posibilă x, ca și orice altă valoare, cu probabilitate nulă. Totuși, probabilitatea ca $\tilde{x}$ să ia valori în intervalul $[x, x+\mathrm{d}x]$ este nenulă dacă x este o valoare posibilă; dacă x nu este valoare posibilă, această afirmație poate să nu mai fie adevărată.

Am utilizat aceeași notație pentru funcția de masă și pentru densitatea de probabilitate deoarece va fi totdeauna clar din context dacă ne referim la o variabilă discretă sau la una continuă.

Scanned with OKEN Scanner

Exemplul 1. Să considerăm variabila aleatoare discretă $\tilde{x}$ care ia valorile $x_1 = -1$, $x_2 = 1$, $x_3 = 3$, respectiv cu probabilitățile $p_1 = 1/6$, $p_2 = 3/6$, $p_3 = 2/6$. Funcția de masă $f(x)$ a acestei variabile aleatoare este

$$f(x) = \begin{cases} 0 & \text{pentru } x \notin \{-1, 1, 3\}, \\ 1/6 & \text{pentru } x = -1, \\ 3/6 & \text{pentru } x = 1, \\ 2/6 & \text{pentru } x = 3, \end{cases}$$

iar funcția de repartiție F este

$$F(x) = \begin{cases} 0 & \text{pentru } x \leqslant -1, \\ 1/6 & \text{pentru } -1 < x \leqslant 1, \\ 4/6 & \text{pentru } 1 < x \leqslant 3, \\ 1 & \text{pentru } x > 3. \end{cases}$$

Graficele funcțiilor f și F sînt prezentate respectiv în fig. 4.2.2 (a) și 4.2.2 (b).

Exemplul 2. Să considerăm variabila aleatoare continuă $\tilde{x}$ a cărei densitate de probabilitate este

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

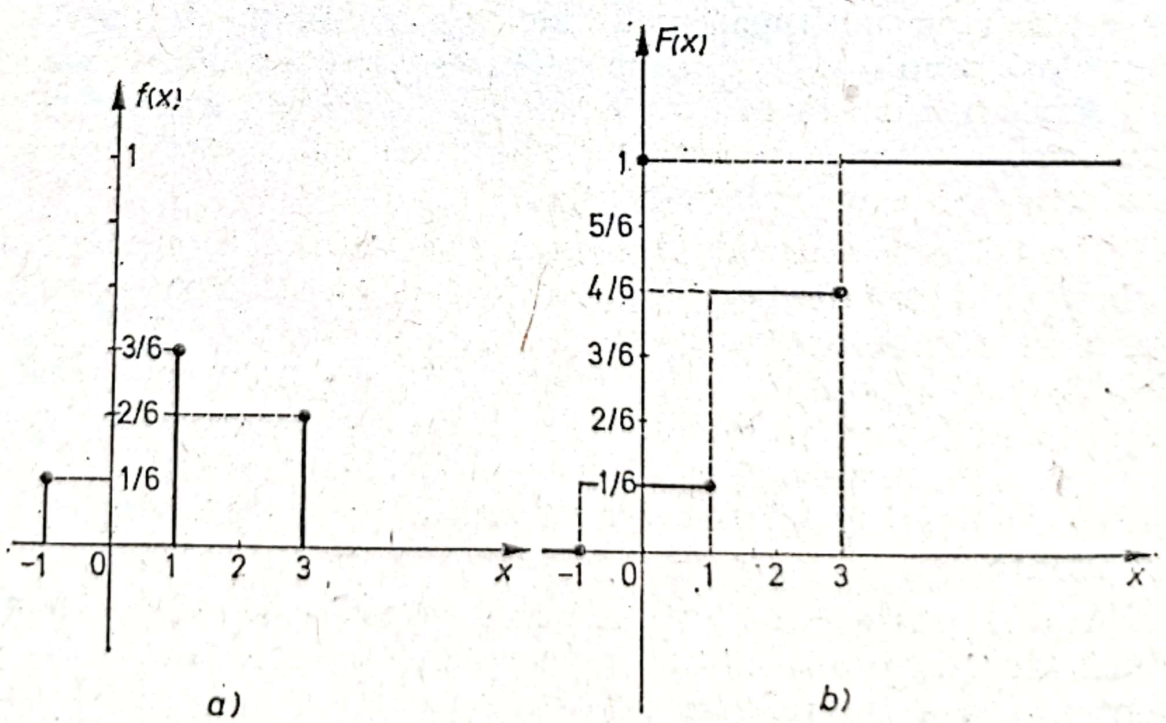


Fig. 4.2.2

Faptul că $f(x) \geqslant 0$ pentru orice $x \in R$ este evident, iar faptul că

$$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1$$

rezultă ușor calculînd integrala *

$$\int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2}\right)\mathrm{d}x = \sqrt{2\pi}.$$

Urmează deci că f verifică proprietățile (DP_1) și (DP_2) care caracterizează o densitate de probabilitate și este prin urmare o densitate de probabilitate. O variabilă aleatoare continuă avînd această densitate de probabilitate este numită și *variabilă aleatoare normală* și joacă un rol important în teoria probabilităților; vom reveni pe larg asupra acestei variabile aleatoare în capitolul 12.

Funcția de repartiție a acestei variabile aleatoare se notează prin $\Phi(x)$, este dată de relația

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right)\mathrm{d}t$$

și poartă numele de *funcția lui Laplace-Gauss;* tabele conținînd valorile acestei funcții sînt date în majoritatea cărților consacrate teoriei probabilităților.

Graficul densității de probabilitate este dat în fig. 4.2.4, iar graficul funcției de repartiție în fig. 4.2.3.

În general, dacă F este funcția de repartiție, iar f este densitatea de probabilitate a unei variabile aleatoare continue $\tilde{x}$, putem scrie

$$P(x \leqslant \tilde{x} \leqslant y) = F(y) - F(x) = \int_{x}^{y} f(t)\,\mathrm{d}t,$$

cum se vede ușor, utilizînd proprietățile integralei. Rezultă că probabilitatea ca $\tilde{x}$ să ia valori în intervalul $[x, y]$ se poate

* Vezi Anexa, § A.1.

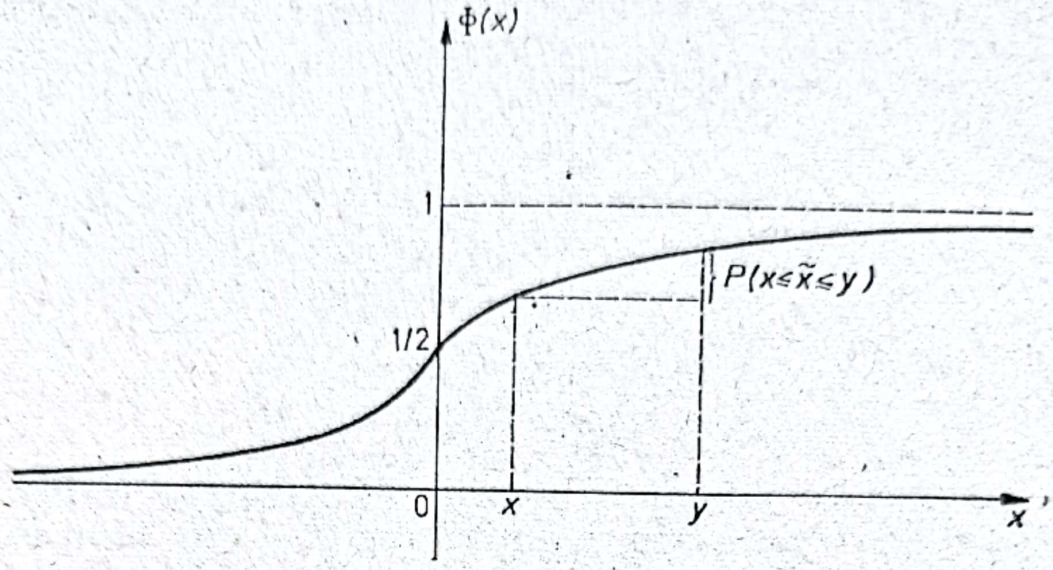


Fig. 4.2.3

exprima fie cu ajutorul funcției de repartiție (vezi fig. 4.2.3), fie cu ajutorul densității de probabilitate (vezi fig. 4.2.4); în ultimul caz, probabilitatea căutată este egală cu aria suprafeței delimitate de graficul densității de probabilitate, axa Ox și dreptele paralele la axa Oy care trec prin extremitățile intervalului $[x, y]$.

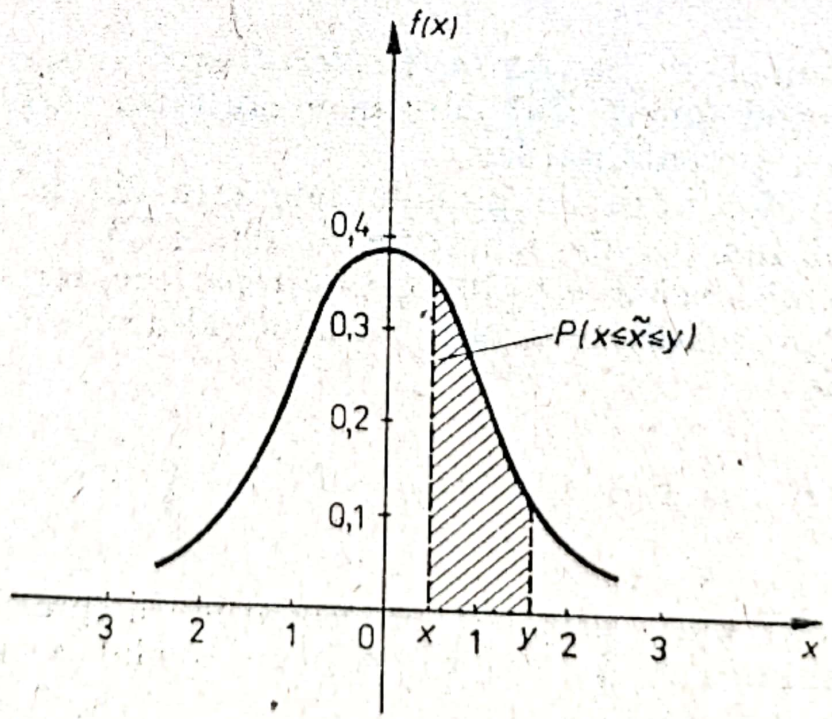


Fig. 4.2.4

4.3. VECTORI ALEATORI. FUNCȚII DE REPARTIȚIE MULTIDIMENSIONALE

Să considerăm un cîmp borelian de probabilitate $(E, \mathcal{K}, P)$ și cîmpul borelian real n-dimensional $(R^n, \mathcal{B}^n)$.

Funcția măsurabilă $\tilde{\mathbf{x}} : E \to R^n$ este numită *variabilă aleatoare n-dimensională* sau *vector aleator*. Pentru fiecare $\omega \in E$, $\tilde{\mathbf{x}}(\omega)$ este un vector cu n componente reale, pe care le vom nota prin $\tilde{x}_1(\omega), \ldots, \tilde{x}_n(\omega)$. Se poate arăta că

$$\tilde{\mathbf{x}}^{-1}(\mathcal{B}^n) = \mathcal{B}(\tilde{x}_1, \ldots, \tilde{x}_n),$$

de unde rezultă că $\tilde{\mathbf{x}}$ este un vector aleator dacă și numai dacă componentele sale $\tilde{x}_1, \ldots, \tilde{x}_n$ sînt variabile aleatoare. În acord cu convenția făcută pentru variabile aleatoare, vom nota $\tilde{\mathbf{x}}(\omega)$ prin $\mathbf{x}$ și $\tilde{x}_1(\omega), \ldots, \tilde{x}_n(\omega)$ respectiv prin $x_1, \ldots, x_n$.

Pentru orice vector $(a_1, \ldots, a_n) \in R^n$ definim *intervalul n-dimensional* $I(a_1, \ldots, a_n)$ în modul următor:

$$I(a_1, \ldots, a_n) = \{(x_1, \ldots, x_n) : x_1 < a_1, \ldots, x_n < a_n\}.$$

Evident, $I(a_1, \ldots, a_n) \in \mathcal{B}^n$.

Definiția 4.3.1. *Funcția* $F : R^n \to R$ *definită de*

$$F(x_1, \ldots, x_n) = P(\{\omega : \tilde{x}_1 < x_1, \ldots, \tilde{x}_n < x_n\}) \qquad (4.3.1)$$

este numită funcția de repartiție a vectorului aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)$ *sau funcția de repartiție multidimensională (mixtă) a variabilelor aleatoare* $\tilde{x}_1, \ldots, \tilde{x}_n$.

Din această definiție rezultă că

$$F(x_1, \ldots, x_n) = P(\tilde{\mathbf{x}}^{-1}(I(x_1, \ldots, x_n))) = P\tilde{\mathbf{x}}(I(x_1, \ldots, x_n)).$$

Funcția de repartiție multidimensională are proprietăți analoage celor demonstrate pentru funcția de repartiție unidimensională. Mai precis, se poate arăta că:

(a). $F(x_1, \ldots, x_n) = 0$ *dacă există cel puțin un indice* i, $1 \leqslant i \leqslant n$, *pentru care* $x_i = -\infty$. *De asemenea, avem* $F(+\infty, \ldots, +\infty) = 1$;

(b). *Funcția* F *este crescătoare în fiecare argument;*

(c). *Funcția* F *este continuă la stînga în fiecare argument.*

Vom căuta acum să obținem o formulă analoagă cu formula (4.2.3), adică vom calcula cu ajutorul funcției de repartiție $F(x_1, ..., x_n)$ probabilitatea ca $(\tilde{x}_1, ..., \tilde{x}_n)$ să aparțină intervalului n-dimensional definit de condițiile

$$\{(x_1, ..., x_n) : a_1 \leqslant x_1 < b_1, ..., a_n \leqslant x_n < b_n\},$$

unde $a_k < b_k$, $1 \leqslant k \leqslant n$.

Se poate stabili ușor formula

$$P(a_1 \leqslant \tilde{x}_1 < b_1, ..., a_n \leqslant \tilde{x}_n < b_n) = \tag{4.3.2}$$

$$= F(b_1, ..., b_n) - \sum_{i=1}^{n} p_i + \sum_{1 \leqslant i < j \leqslant n} p_{ij} + ... + (-1)^n F(a_1, ..., a_n),$$

unde $p_{ij...k}$ reprezintă valoarea funcției de repartiție $F(x_1, ..., x_n)$ pentru $x_i = a_i$, $x_j = a_j$, ..., $x_k = a_k$, restul variabilelor x_s fiind egale cu b_s.

De exemplu, pentru $n = 2$, formula devine

$$P(a_1 \leqslant \tilde{x}_1 < b_1, a_2 \leqslant \tilde{x}_2 < b_2) =$$

$$= F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2). \tag{4.3.3}$$

Se poate vedea ușor că proprietățile *(a)*, *(b)*, *(c)* de mai sus nu caracterizează funcțiile de repartiție multidimensionale. Cu alte cuvinte, dacă F este o funcție care verifică aceste proprietăți, nu rezultă neapărat că ea este funcția de repartiție a unui anumit vector aleator. Se poate arăta însă că dacă, în plus, pentru orice interval n-dimensional definit de numerele reale $a_i < b_i$, $1 \leqslant i \leqslant n$, membrul drept al relației (4.3.2) este nenegativ, atunci F este într-adevăr funcția de repartiție a unui vector aleator n-dimensional.

O altă problemă interesantă este calculul funcțiilor de repartiție ale variabilelor aleatoare $\tilde{x}_1, ..., \tilde{x}_n$ cu ajutorul funcției de repartiție $F(x_1, ..., x_n)$ a vectorului aleator $(\tilde{x}_1, ..., \tilde{x}_n)$. Funcția de repartiție a variabilei aleatoare $\tilde{x}_i$ va fi notată prin $F_i(x_i)$, $1 \leqslant i \leqslant n$, și va fi numită și *funcție de repartiție marginală*. Se poate vedea ușor că

$$F_i(x_i) = F(+\infty, ..., x_i, ..., +\infty), \quad 1 \leqslant i \leqslant n. \tag{4.3.4}$$

Cu alte cuvinte, $F_i(x_i)$ se obține din $F(x_1, ..., x_n)$ punînd $x_j = +\infty$ pentru $j \neq i$.

Scanned with OKEN Scanner

În mod analog se poate calcula funcția de repartiție $F_{i_1, \ldots, i_k}$ a vectorului aleator $(\widetilde{x}_{i_1}, \ldots, \widetilde{x}_{i_k})$ format cu k $(1 \leqslant k \leqslant n)$ componente ale vectorului $(\widetilde{x}_1, \ldots, \widetilde{x}_n)$, a cărui funcție de repartiție $F(x_1, \ldots, x_n)$ este cunoscută:

$$F_{i_1, \ldots, i_k}(x_{i_1}, \ldots, x_{i_k}) = F(y_1, \ldots, y_n), \tag{4.3.5}$$

unde $y_j = x_j$ pentru $j = i_s$, $1 \leqslant s \leqslant k$, și $y_j = +\infty$ în celelalte cazuri.

Dacă există o funcție măsurabilă $f : R^n \to R$ astfel încît funcția de repartiție multidimensională $F(x_1, \ldots, x_n)$ să se exprime sub forma

$$F(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_n} f(t_1, \ldots, t_n)\, dt_1 \ldots dt_n, \tag{4.3.6}$$

atunci funcția f este numită *densitatea de probabilitate* a vectorului aleator $\widetilde{\mathbf{x}} = (\widetilde{x}_1, \ldots, \widetilde{x}_n)$ sau *densitatea de probabilitate multidimensională* (comună). Din proprietățile funcției de repartiție rezultă că densitatea de probabilitate f are proprietățile următoare:

(DP_1) $f(x_1, \ldots, x_n) \geqslant 0$ *pentru orice* $(x_1, \ldots, x_n) \in R^n$;

(DP_2) $\displaystyle\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n)\, dx_1 \ldots dx_n = 1.$

În plus, avem

$$\frac{\partial^n F(x_1, \ldots, x_n)}{\partial x_1 \ldots \partial x_n} = f(x_1, \ldots, x_n), \tag{4.3.7}$$

aproape peste tot în raport cu măsura Lebesgue pe R^n.

De asemenea, se poate arăta că

$$P(x_1 \leqslant \widetilde{x}_1 \leqslant x_1 + dx_1, \ldots, x_n \leqslant \widetilde{x}_n \leqslant x_n + dx_n) \approx$$
$$\approx f(x_1, \ldots, x_n)\, dx_1 \ldots dx_n,$$

formulă care ne permite interpretarea probabilistă a densității de probabilitate multidimensională.

Este clar din cele de mai sus că, dacă vectorul aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, ..., \tilde{x}_n)$ are o densitate de probabilitate f, atunci componentele sale au și ele densități de probabilitate și deci sînt variabile aleatoare continue. Mai precis, densitatea de probabilitate a variabilei aleatoare $\tilde{x}_i$, pe care o notăm prin $f_i(x_i)$, se determină din formula

$$f_i(x_i) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f(x_1, ..., x_n)\, dx_1 \dots dx_{i-1}\, dx_{i+1} \dots dx_n,$$

adică prin integrarea densității de probabilitate multidimensionale în raport cu toate variabilele x_j, $j \neq i$.

În analogie cu cazul unidimensional, putem considera și vectori aleatori discreți. Vom spune că $\tilde{\mathbf{x}} = (\tilde{x}_1, ..., \tilde{x}_n)$ este un *vector aleator discret* dacă mulțimea valorilor sale $(\mathbf{x}^i)_{i \in I}$ este o mulțime cel mult numărabilă. Putem defini și în acest caz o funcție de masă. Funcția $f : R^n \to R$ dată de relația

$$f(\mathbf{x}) = \begin{cases} 0 \text{ dacă } \mathbf{x} \notin (\mathbf{x}^i)_{i \in I}, \\ P(\tilde{\mathbf{x}} = \mathbf{x}^i) \text{ dacă } \mathbf{x} = \mathbf{x}^i,\ i \notin I, \end{cases} \tag{4.3.8}$$

este numită *funcția de masă* a vectorului aleator $\tilde{\mathbf{x}}$.

Este clar că funcția de repartiție F a vectorului aleator discret $\tilde{\mathbf{x}}$ este dată de relația

$$F(\mathbf{x}) = \sum_{\{i : \mathbf{x}^i < \mathbf{x}\}} f(\mathbf{x}^i). \tag{4.3.9}$$

În plus, funcția de masă f definită de (4.3.8) are proprietățile:

(FM_1) $\mathrm{f}(\mathbf{x}) \geqslant 0$ *pentru orice* $\mathbf{x} \in \mathrm{R}^n$;

(FM_2) $\mathrm{f}(\mathbf{x}) = 0$ *pentru orice* $\mathbf{x} \notin (\mathbf{x}^i)_{i \in I}$;

(FM_3) $\sum_{i \in I} \mathrm{f}(\mathbf{x}^i) = 1$.

Evident, funcția de masă a unei variabile aleatoare $\tilde{x}_i$, componentă a vectorului aleator discret $\tilde{\mathbf{x}} = (\tilde{x}_1, ..., \tilde{x}_n)$, pe care o vom nota prin $f_i(x_i)$, este dată de relația

$$f_i(x_i) = \sum_{x_1, \dots x_{i-1}, x_{i+1}, \dots, x_n} f(x_1, ..., x_{i-1}, x_i, x_{i+1}, ..., x_n).$$

În cazul multidimensional putem considera și vectori aleatori „hibrizi", avînd unele componente continue, iar altele

discrete. Analogia dintre proprietățile densității de probabilitate și ale funcției de masă care au fost prezentate mai sus ne permite să introducem pentru acești vectori noțiunea de *funcție elementară de repartiție* sau *repartiție elementară*. Vom spune că f este repartiția elementară a vectorului aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, ..., \tilde{x}_n)$ dacă funcția sa de repartiție F se exprimă sub forma

$$F(x_1, ..., x_n) = \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_n} f(t_1, ..., t_n)\, dt_1 ... dt_n \quad (4.3.10)$$

și are proprietățile următoare:

(RE_1) $f(x_1, ..., x_n) \geqslant 0$ *pentru orice* $(x_1, ..., x_n) \in R^n$;

$$(RE_2) \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(x_1, ..., x_n)\, dx_1 ... dx_n = 1. \quad (4.3.11)$$

În (4.3.10) cît și în (4.3.11), integralele se înlocuiesc cu sume dacă componenta respectivă a vectorului aleator este discretă. Evident, densitatea de probabilitate și funcția de masă multidimensionale sînt cazuri particulare ale repartiției elementare. Se poate arăta însă că repartiția elementară marginală se află după aceeași regulă pe care am pus-o în evidență mai sus pentru cele două cazuri particulare.

Exemplul 1. Să considerăm vectorul aleator discret $\tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2)$ a cărui funcție de masă este

$$f(x_1, x_2) = \begin{cases} 1/10 & \text{pentru } (x_1, x_2) = (1, -1), \\ 3/10 & \text{pentru } (x_1, x_2) = (2, -1), \\ 1/10 & \text{pentru } (x_1, x_2) = (1, 0), \\ 5/10 & \text{pentru } (x_1, x_2) = (2, 0), \\ 0 & \text{în celelalte cazuri.} \end{cases}$$

Se constată fără dificultate că această funcție este într-adevăr o funcție de masă. Funcția de masă marginală pentru $\tilde{x}_1$ este

$$f_1(x_1) = \sum_{x_2} f(x_1, x_2) = \begin{cases} 2/10 & \text{pentru } x_1 = 1, \\ 8/10 & \text{pentru } x_1 = 2, \\ 0 & \text{în celelalte cazuri.} \end{cases}$$

Scanned with OKEN Scanner

Analog, funcția de masă marginală a lui $\tilde{x}_2$ este

$$f_2(x_2) = \sum_{x_1} f(x_1, x_2) = \begin{cases} 4/10 & \text{pentru } x_2 = -1, \\ 6/10 & \text{pentru } x_2 = 0, \\ 0 & \text{în celelalte cazuri.} \end{cases}$$

Exemplul 2. Să considerăm vectorul aleator continuu $\tilde{x} = (\tilde{x}_1, \tilde{x}_2)$ a cărui densitate de probabilitate este

$$f(x_1, x_2) = \begin{cases} 1/6 & \text{pentru } 0 \leqslant x_1 \leqslant 2 \text{ și } 0 \leqslant x_2 \leqslant 3, \\ 0 & \text{în celelalte cazuri.} \end{cases}$$

Deoarece

$$\int_0^2 \int_0^3 f(x_1, x_2)\, dx_1\, dx_2 = 1,$$

rezultă că f este într-adevăr o densitate de probabilitate.

Densitatea de probabilitate marginală pentru $\tilde{x}_1$ este

$$f_1(x_1) = \int_0^3 f(x_1, x_2)\, dx_2 = \begin{cases} 1/2 & \text{pentru } 0 \leqslant x_1 \leqslant 2, \\ 0 & \text{în caz contrar,} \end{cases}$$

iar densitatea marginală pentru $\tilde{x}_2$ este

$$f_2(x_2) = \int_0^2 f(x_1, x_2)\, dx_1 = \begin{cases} 1/3 & \text{pentru } 0 \leqslant x_2 \leqslant 3, \\ 0 & \text{în caz contrar.} \end{cases}$$

În sfîrșit, funcția de repartiție $F(x_1, x_2)$ se calculează din formula cunoscută și este

$$F(x_1, x_2) = \begin{cases} 0 & \text{dacă } x_1 \leqslant 0 \text{ sau } x_2 \leqslant 0, \\ \dfrac{x_1 x_2}{6} & \text{dacă } 0 < x_1 \leqslant 2 \text{ și } 0 < x_2 \leqslant 3, \\ x_1/2 & \text{dacă } 0 < x_1 \leqslant 2 \text{ și } x_2 > 3, \\ x_2/3 & \text{dacă } 0 < x_2 \leqslant 3 \text{ și } x_1 > 2, \\ 1 & \text{dacă } x_2 > 3 \text{ și } x_1 > 2. \end{cases}$$

Scanned with OKEN Scanner

Exemplul 3. Să considerăm vectorul aleator $\tilde{x} = (\tilde{x}_1, \tilde{x}_2)$, a cărui repartiție elementară este dată prin

$$f(x_1, x_2) = \begin{cases} 1/3 \text{ dacă } 0 \leqslant x_1 \leqslant 1 \text{ și } x_2 = 0, \\ 2/3 \text{ dacă } 0 \leqslant x_1 \leqslant 1 \text{ și } x_2 = 1, \\ 0 \text{ în celelalte cazuri.} \end{cases}$$

Se constată mai întîi că f este într-adevăr o repartiție elementară, deoarece

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1\, dx_2 = \int_0^1 \left(\sum_{x_2} f(x_1, x_2) \right) dx_1 =$$

$$= \int_0^1 (1/3 + 2/3)\, dx_1 = 1.$$

Prima componentă a vectorului aleator este o variabilă aleatoare continuă, a cărei densitate de probabilitate marginală este:

$$f_1(x_1) = \sum_{x_2} f(x_1, x_2) = \begin{cases} 1 \text{ dacă } 0 \leqslant x_1 \leqslant 1, \\ 0 \text{ în caz contrar.} \end{cases}$$

A doua componentă este o variabilă aleatoare discretă a cărei funcție de masă este

$$f_2(x_2) = \int_0^1 f(x_1, x_2)\, dx_1 = \begin{cases} 1/3 \text{ pentru } x_2 = 0, \\ 2/3 \text{ pentru } x_2 = 1, \\ 0 \text{ în celelalte cazuri.} \end{cases}$$

4.4. VARIABILE ALEATOARE INDEPENDENTE

Fie $(E, \mathcal{K}, P)$ un cîmp borelian de probabilitate și fie $\tilde{x}_1, \ldots, \tilde{x}_n$ variabile aleatoare pe acest cîmp.

Definiția 4.4.1. *Variabilele aleatoare* $\tilde{x}_1, \ldots, \tilde{x}_n$ *se numesc independente dacă, pentru orice numere reale* $x_1, \ldots, x_n$, *evenimentele* $\tilde{x}_1^{-1}((-\infty, x_1)), \ldots, \tilde{x}_n^{-1}((-\infty, x_n))$ *sînt independente.*

Cu alte cuvinte, oricare ar fi $k \leqslant n$ și oricare ar fi numerele reale $x_{i_1}, \ldots, x_{i_k}$, trebuie să avem

$$P(\widetilde{x}_{i_1} < x_{i_1}, \ldots, \widetilde{x}_{i_k} < x_{i_k}) = P(\widetilde{x}_{i_1} < x_{i_1}) \ldots P(\widetilde{x}_{i_n} < x_{i_n}). \quad (4.4.1)$$

În particular, rezultă că

$$P(\widetilde{x}_1 < x_1, \ldots, \widetilde{x}_n < x_n) = P(\widetilde{x}_1 < x_1) \ldots P(\widetilde{x}_n < x_n)$$

sau

$$F(x_1, \ldots, x_n) = F_1(x_1) \ldots F_n(x_n), \quad (4.4.2)$$

unde $F_i(x_i)$, $1 \leqslant i \leqslant n$, sînt funcțiile de repartiție marginale ale variabilelor aleatoare $\widetilde{x}_i$, $1 \leqslant i \leqslant n$, iar $F(x_1, \ldots, x_n)$ este funcția de repartiție a vectorului aleator $(\widetilde{x}_1, \ldots, \widetilde{x}_n)$.

Din cele de mai sus rezultă că (4.4.2) este o condiție necesară de independență pentru variabilele aleatoare $\widetilde{x}_1, \ldots, \widetilde{x}_n$. Putem arăta că această condiție este și suficientă. Într-adevăr, punînd în (4.4.2) $x_j = +\infty$ pentru $j \notin \{i_1, \ldots, i_k\}$ se obține imediat (4.4.1), oricare ar fi $i_1 < i_2 < \ldots < i_k$ și $k \leqslant n$.

Se poate de asemenea vedea că, dacă variabilele aleatoare independente $\widetilde{x}_1, \ldots, \widetilde{x}_n$ au repartițiile elementare $f_1(x_1), \ldots, f_n(x_n)$, atunci repartiția elementară a vectorului $(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ este dată de relația

$$f(x_1, \ldots, x_n) = f_1(x_1) \ldots f_n(x_n). \quad (4.4.3)$$

Noțiunea de independență se poate extinde pe calea obișnuită la familii oarecare de variabile aleatoare.

Definiția 4.4.2. *Variabilele aleatoare* $(\widetilde{x}_i)_{i \in I}$ *se numesc independente dacă, pentru orice submulțime finită* J *a lui* I, *variabilele aleatoare* $(\widetilde{x}_i)_{i \in J}$ *sînt independente.*

Cu alte cuvinte, variabilele aleatoare $(\widetilde{x}_i)_{i \in I}$ sînt independente dacă pentru orice număr natural n și orice indici $i_1, \ldots, i_n \in I$, avem

$$F(x_{i_1}, \ldots, x_{i_n}) = F_{i_1}(x_{i_1}) \ldots F_{i_n}(x_{i_n}),$$

pentru orice numere reale $x_{i_1}, \ldots, x_{i_n}$, unde $F(x_{i_1}, \ldots, x_{i_n})$ este funcția de repartiție a vectorului aleator $(\widetilde{x}_{i_1}, \ldots, \widetilde{x}_{i_n})$, iar $F_{i_j}(x_{i_j})$, $1 \leqslant j \leqslant n$, sînt funcțiile de repartiție marginale ale variabilelor aleatoare $\widetilde{x}_{i_1}, \ldots, \widetilde{x}_{i_n}$.

Teorema 4.4.1. *Variabilele aleatoare* $(\widetilde{x}_i)_{i \in I}$ *sînt independente dacă și numai dacă pentru orice parte finită* $J \subset I$ *și orice familie finită* $(A_i)_{i \in J} \subset \mathcal{B}$, *avem*

$$P\left(\bigcap_{i \in J} \widetilde{x}_i^{-1}(A_i)\right) = \prod_{i \in J} P(\widetilde{x}_i^{-1}(A_i)). \quad (4.4.4)$$

Demonstrație. Suficiența condiției rezultă imediat dacă luăm $A_i = (-\infty, x_i)$, $i \in J$.

Pentru a arăta necesitatea condițiilor este suficient să considerăm cazul a două variabile aleatoare $\tilde{x}_1$ și $\tilde{x}_2$, cazul unui număr finit arbitrar rezultînd imediat.

Pentru orice $x \in R$ considerăm familia $\mathfrak{M}(x)$ a mulțimilor $B \subset R$ cu proprietatea că $x^{-1}_2(B) \in \mathfrak{K}$ și

$$P(\tilde{x}_1^{-1}((-\infty, x)) \cap \tilde{x}_2^{-1}(B)) = P(\tilde{x}_1^{-1}((-\infty, x)) P(\tilde{x}_2^{-1}(B)). \quad (4.4.5)$$

Este clar că $\mathfrak{M}(x)$ este nevidă și conține intervalele de forma $(-\infty, y)$ cu $y \in R$. Se poate de asemenea vedea ușor că intervalele de forma $[y, z)$ aparțin familiei $\mathfrak{M}(x)$. Într-adevăr, deoarece $[y, z) = (-\infty, z) - (-\infty, y)$, rezultă că putem scrie

$$\begin{aligned} & P(\tilde{x}_1^{-1}((-\infty, x)) \cap \tilde{x}_2^{-1}([y, z))) = \\ & = P(\tilde{x}_1^{-1}((-\infty, \tilde{x})) \cap \tilde{x}_2^{-1}((-\infty, z))) - \\ & - P(\tilde{x}_1^{-1}((-\infty, x)) \cap \tilde{x}_2^{-1}((-\infty, y))) = \\ & = P(\tilde{x}_1^{-1}((-\infty, x)) P(\tilde{x}_2^{-1}((-\infty, z))) - \\ & - P(\tilde{x}_1^{-1}((-\infty, x))) P(\tilde{x}_2^{-1}((-\infty, y))) = \\ & = P(\tilde{x}_1^{-1}((-\infty, x)) P(\tilde{x}_2^{-1}([y, z))) \end{aligned}$$

și deci $[y, z) \in \mathfrak{M}(x)$.

Dacă acum $B \in \mathfrak{B}$, rezultă că B se poate reprezenta ca o reuniune cel mult numărabilă de intervale disjuncte de forma $[y, z)$, adică $B = \bigcup_{n \in N} [y_n, z_n)$. Rezultă atunci

$$\begin{aligned} & P((\tilde{x}_1^{-1}((-\infty, \tilde{x})) \cap \tilde{x}_2^{-1}\left(\bigcup_{n \in N} [y_n, z_n)\right)) = \\ & = \sum_{n \in N} P(\tilde{x}_1^{-1}((-\infty, x)) \cap \tilde{x}_2^{-1}([y_n, z_n))) = \\ & = \sum_{n \in N} P(\tilde{x}_1^{-1}((-\infty, x)) P(\tilde{x}_2^{-1}([y_n, z_n))) = \\ & = P(\tilde{x}_1^{-1}((-\infty, x))) P(\tilde{x}_2^{-1}(B)) \end{aligned}$$

și deci $B \in \mathfrak{M}(x)$. Prin urmare relația (4.4.5) este valabilă pentru orice $B \in \mathfrak{B}$ și orice $x \in R$.

Scanned with OKEN Scanner

Fie acum $B \in \mathcal{B}$ fixat și fie $\mathcal{R}(B)$ formată cu toate mulțimile $A \subset R$ cu proprietatea că $\tilde{x}_1^{-1}(A) \in \mathcal{K}$ și

$$P(\tilde{x}_1^{-1}(A) \cap \tilde{x}_2^{-1}(B)) = P(\tilde{x}_1^{-1}(A))\ P(\tilde{x}_2^{-1}(B)). \qquad (4.4.6)$$

Este clar că $\mathcal{R}(B)$ este nevidă și conține toate intervalele de forma $(-\infty, x)$ cu $x \in R$. La fel ca mai sus se arată că orice $A \in \mathcal{B}$ aparține lui $\mathcal{R}(B)$ și deci (4.4.6) este verificată de orice $A \in \mathcal{B}$. Cum $B \in \mathcal{B}$ este arbitrar, rezultă că (4.4.6) este satisfăcută de orice A, $B \in \mathcal{B}$ și teorema este demonstrată.

Din această teoremă rezultă că independența familiei $(\tilde{x}_i)_{i \in I}$ de variabile aleatoare poate fi definită echivalent prin condiția (4.4.4).

Un rezultat util este următorul:

Teorema 4.4.2. *Dacă* $(\tilde{x}_i)_{i \in I}$ *sînt variabile aleatoare independente, iar* $(\tilde{y}_i)_{i \in I}$ *sînt funcții de variabilă reală măsurabile, atunci variabilele aleatoare* $(\tilde{y}_i \circ \tilde{x}_i)_{i \in I}$ *sînt de asemenea independente.*

Demonstrație. Fie n natural, fie $i_1, \ldots, i_n \in I$ și fie $x_{i_1}, \ldots, x_{i_n}$ numere reale arbitrare. Dacă notăm $\tilde{z}_i = \tilde{y}_i \circ \tilde{x}_i$, utilizînd un rezultat precedent, obținem

$$P\left(\bigcap_{j=1}^{n} \tilde{z}_{i_j}^{-1}((-\infty, x_{i_j}))\right) = P\left(\bigcap_{j=1}^{n} \tilde{x}_{i_j}^{-1}(\tilde{y}_{i_j}^{-1}((-\infty, x_{i_j})))\right) =$$

$$= \prod_{j=1}^{n} P(\tilde{x}_{i_j}^{-1}(\tilde{y}_{i_j}^{-1}(-\infty, x_{i_j}))) = \prod_{j=1}^{n} P(\tilde{z}_{i_j}^{-1}((-\infty, x_{i_j}))),$$

deoarece $\tilde{y}_{i_j}^{-1}((-\infty, x_{i_j})) \in \mathcal{B}$.

După cum se știe, unei variabile aleatoare i se asociază un corp borelian, numit corpul borelian generat de acea variabilă aleatoare. Dacă se consideră o familie $(\tilde{x}_i)_{i \in I}$ de variabile aleatoare independente, atunci îi putem asocia familia corpurilor boreliene generate $(\mathcal{B}(\tilde{x}_i))_{i \in I}$. Este natural să ne întrebăm dacă $(\mathcal{B}(\tilde{x}_i))_{i \in I}$ sînt corpuri boreliene independente. Avem în acest sens rezultatul următor.

Scanned with OKEN Scanner

Teorema 4.4.3. *Variabilele aleatoare* $(\tilde{x}_i)_{i \in I}$ *sînt independente dacă și numai dacă corpurile boreliene generate* $(\mathcal{B}(\tilde{x}_i))_{i \in I}$ *sînt independente.*

Demonstrație. Rezultatul se obține imediat din teorema 4.4.1.

Din această propoziție rezultă că putem defini direct independența variabilelor aleatoare prin independența corpurilor boreliene generate. Putem deci vorbi de independența unei familii formate din evenimente, desfaceri, variabile aleatoare și corpuri boreliene, înțelegînd prin aceasta independența corpurilor boreliene generate asociate acestei familii.

Scanned with OKEN Scanner

Capitolul 5

VALORI NUMERICE ASOCIATE VARIABILELOR ALEATOARE

Funcția de repartiție a unei variabile aleatoare constituie o descriere completă a fenomenului aleator considerat. În multe probleme practice nu avem nevoie de o descriere atît de detaliată, putîndu-ne mulțumi cu informații parțiale. În acest scop, unei variabile aleatoare i se asociază anumite valori numerice, care ne permit să ne formăm o imagine aproximativă despre fenomenul aleator considerat, mult mai puțin precisă decît cea oferită de funcția de repartiție. Printre valorile numerice care se asociază unei variabile aleatoare, cele mai importante sînt: media, dispersia, momentele de diferite ordine etc.

5.1. INTEGRALA LEBESGUE

Să considerăm un cîmp borelian de probabilitate $(E, \mathfrak{K}, P)$ și o variabilă aleatoare $\tilde{x}$ pe acest cîmp care ia numai valori nenegative (eventual și valoarea $+\infty$). Acestei variabile aleatoare i se asociază un număr real notat prin

$$\int_E \tilde{x}(\omega)\, dP(\omega)$$

și numit *integrala Lebesgue** a lui $\tilde{x}$ în raport cu P, caracterizat

* Uneori vom scrie integrala Lebesgue sub forma $\int_E \tilde{x}(\omega)\, P(d\omega)$, $\int_E \tilde{x}\, dP$ sau încă $\int \tilde{x}\, dP$, cînd nu există pericol de confuzie.

de proprietățile următoare:

(*i*) *Dacă* $\tilde{x}$ *și* $\tilde{y}$ *sînt variabile aleatoare nenegative, atunci:*

$$\int_E [a\tilde{x}(\omega) + b\tilde{y}(\omega)]\, dP(\omega) = a\int_E \tilde{x}(\omega)\, dP(\omega) + b\int_E \tilde{y}(\omega)\, dP(\omega), \tag{5.1.1}$$

unde a *și* b *sînt numere reale nenegative;*

(*ii*) *Dacă* $A \in \mathfrak{K}$, *atunci*

$$\int_E \chi_A(\omega)\, dP(\omega) = P(A); \tag{5.1.2}$$

(*iii*) *Dacă* $(\tilde{x}_n)$ *este un șir crescător de variabile aleatoare nenegative, atunci*

$$\int_E \left(\lim_n \tilde{x}_n(\omega)\right) dP(\omega) = \lim_n \int_E \tilde{x}_n(\omega)\, dP(\omega). \tag{5.1.3}$$

Construcția integralei $\int \tilde{x}\, dP$ este făcută mai întîi pentru o variabilă aleatoare simplă, de forma

$$\tilde{x} = x_1\chi_{A_1} + \ldots + x_n\chi_{A_n}$$

cu $A_1, \ldots, A_n \in \mathfrak{K}$. Din (*i*) și (*ii*) rezultă că în acest caz trebuie să avem

$$\int_E \tilde{x}\, dP = \sum_{i=1}^{n} x_i P(A_i). \tag{5.1.4}$$

Construcția continuă apoi pentru variabile aleatoare nenegative arbitrare, utilizînd proprietatea (*iii*) și faptul că orice variabilă aleatoare nenegativă este limita unui șir crescător de variabile aleatoare nenegative simple (vezi § 4.1, teorema 4.1.2).

Vom spune că o variabilă aleatoare $\tilde{x}$ pe $(E, \mathfrak{K}, P)$ este *integrabilă* sau *sumabilă* în raport cu P dacă integrala Lebesgue a lui $|\tilde{x}|$ este finită, adică

$$\int_E |\tilde{x}(\omega)|\, dP(\omega) < +\infty. \tag{5.1.5}$$

Scanned with OKEN Scanner

Dacă notăm

$$\tilde{x}^+(\omega) = \max(\tilde{x}(\omega), 0); \quad \tilde{x}^-(\omega) = \max(-\tilde{x}(\omega), 0),$$

atunci se poate vedea ușor că $\tilde{x}$ este integrabilă dacă și numai dacă variabilele aleatoare nenegative $\tilde{x}^+$ și $\tilde{x}^-$ sînt integrabile. Integrala Lebesgue a variabilei aleatoare $\tilde{x}$ integrabile este definită de relația

$$\int_E \tilde{x}(\omega)\, dP(\omega) = \int_E \tilde{x}^+(\omega)\, dP(\omega) - \int_E \tilde{x}^-(\omega)\, dP(\omega). \quad (5.1.6)$$

Mulțimea variabilelor aleatoare integrabile pe $(E, \mathcal{K}, P)$ va fi notată prin $\mathfrak{L}(E, \mathcal{K}, P)$ sau $\mathfrak{L}$, dacă nu există pericol de confuzie.

Este clar că proprietatea (i) se menține și pentru variabile aleatoare integrabile $\tilde{x}$ și $\tilde{y}$ oarecare, numerele a și b putînd fi de asemenea arbitrare. Amintim totodată că pentru variabilele aleatoare nenegative au loc și proprietățile:

(iv) *Dacă* $\tilde{x}_1 \leqslant \tilde{x}_2$ *atunci*

$$\int_E \tilde{x}_1(\omega)\, dP(\omega) \leqslant \int_E \tilde{x}_2(\omega)\, dP(\omega);$$

(v) *Dacă* $(\tilde{x}_n)$ *este un șir descrescător, iar* $\tilde{x}_1$ *este integrabilă, atunci*

$$\int_E \left(\lim_n \tilde{x}_n(\omega)\right) dP(\omega) = \lim_n \int_E \tilde{x}_n(\omega)\, dP(\omega); \quad (5.1.7)$$

(vi) *Dacă* $\tilde{x}$ *este integrabilă, atunci*

$$P(\{\omega : \tilde{x}(\omega) = +\infty\}) = 0. \quad (5.1.8)$$

Rezultatul care urmează arată legătura dintre integrala Lebesgue și integrala Stieltjes.

Teorema 5.1.1. *Fie* $\tilde{x}$ *o variabilă aleatoare pe* $(E, \mathcal{K}, P)$ *și fie* F *funcția sa de repartiție. Dacă* $g : R \to R$ *este măsurabilă,*

iar variabila aleatoare $\tilde{z}(\omega) = g(\tilde{x}(\omega))$ *este integrabilă Lebesgue, atunci*

$$\int_E g(\tilde{x}(\omega))\, dP(\omega) = \int_R g(x)\, dF(x), \tag{5.1.9}$$

unde termenul din membrul drept este integrala Lebesgue-Stieltjes a lui g *în raport cu* F.

Demonstrație. Vom arăta mai întîi că formula (5.1.9) are loc în cazul cînd g este o funcție măsurabilă simplă, adică este de forma

$$g(x) = \sum_{k=1}^{n} x_k \chi_{A_k}(x),$$

unde $A_1, \ldots, A_n \in \mathfrak{B}$, iar $x_1, \ldots, x_n$ sînt valorile luate de g pe mulțimile $A_1, \ldots, A_n$. Deoarece orice mulțime din $\mathfrak{B}$ este o reuniune cel mult numărabilă de intervale disjuncte de forma $[a, b)$, este suficient să arătăm (5.1.9) pentru aplicațiile de forma $g(x) = \chi_{[a, b)}(x)$. Deoarece variabila aleatoare $\tilde{z}$ devine în acest caz

$$\tilde{z}(\omega) = \begin{cases} 1 \text{ dacă } \tilde{x}(\omega) \in [a, b), \\ 0 \text{ dacă } \tilde{x}(\omega) \notin [a, b), \end{cases}$$

rezultă că

$$\int_E \tilde{z}(\omega)\, dP(\omega) = 1 \cdot P(a \leqslant \tilde{x}(\omega) < b) + 0 \cdot P(\tilde{x}(\omega) \notin [a, b)) =$$

$$= P(a \leqslant \tilde{x}(\omega) < b) = F(b) - F(a).$$

Pe de altă parte, avem

$$\int_R \chi_{[a, b)}(x)\, dF(x) = \int_a^b dF(x) = F(b) - F(a),$$

ceea ce demonstrează formula (5.1.9) în acest caz.

Dacă g este o funcție măsurabilă nenegativă, atunci există un șir crescător de funcții măsurabile simple (g_n) convergent

către g (vezi teorema 4.1.2). Rezultă atunci că șirul cu termenul general $\widetilde{z}_n(\omega) = g_n(\widetilde{x}(\omega))$ este crescător și convergent la $\widetilde{z}(\omega)$ și

$$\int_E \widetilde{z}(\omega)\,\mathrm{d}P(\omega) = \lim_n \int_E \widetilde{z}_n(\omega)\,\mathrm{d}P(\omega) = \lim_n \int_R g_n(x)\,\mathrm{d}F(x) =$$

$$= \int_R g(x)\,\mathrm{d}F(x), \tag{5.1.10}$$

ținînd seama de proprietatea (5.1.3) a integralei Lebesgue și de proprietățile integralei Stieltjes. Relația (5.1.10) ne arată că formula (5.1.9) este verificată pentru $g \geqslant 0$.

Pentru a termina demonstrația, să considerăm o funcție măsurabilă $g : R \to R$ oarecare și să punem $g = g^+ - g^-$. Rezultă că variabila aleatoare $\widetilde{z}(\omega) = g(\widetilde{x}(\omega))$ se scrie sub forma $\widetilde{z} = \widetilde{z}^+ - \widetilde{z}^-$, unde

$$\widetilde{z}^+(\omega) = g^+(\widetilde{x}(\omega)), \quad \widetilde{z}^-(\omega) = g^-(\widetilde{x}(\omega)).$$

Putem deci scrie

$$\int_E \widetilde{z}(\omega)\,\mathrm{d}P(\omega) = \int_E \widetilde{z}^+(\omega)\,\mathrm{d}P(\omega) - \int_E \widetilde{z}^-(\omega)\,\mathrm{d}P(\omega) =$$

$$= \int_R g^+(x)\,\mathrm{d}F(x) - \int_R g^-(x)\,\mathrm{d}F(x) = \int_R g(x)\,\mathrm{d}F(x),$$

ceea ce arată că teorema este adevărată pentru orice funcție măsurabilă g.

În cazul cînd $g(x) = x$, formula (5.1.9) devine

$$\int_E \widetilde{x}(\omega)\,\mathrm{d}P(\omega) = \int_R x\,\mathrm{d}F(x). \tag{5.1.11}$$

Teorema precedentă poate fi generalizată astfel. Dacă $\widetilde{x}_1, \ldots, \widetilde{x}_n$ sînt variabile aleatoare pe $(E, \mathfrak{K}, P)$, funcția $g : R^n \to R$

este măsurabilă, iar variabila aleatoare $\tilde{z}(\omega) = g(\tilde{x}_1(\omega), \ldots, \tilde{x}_n(\omega))$ este integrabilă, atunci avem

$$\int_E g(\tilde{x}_1(\omega), \ldots, \tilde{x}_n(\omega))\, dP(\omega) = \int_{R^n} g(x_1, \ldots, x_n)\, dF(x_1, \ldots, x_n), \quad (5.1.12)$$

unde $F(x_1, \ldots, x_n)$ este funcția de repartiție a vectorului aleator $(\tilde{x}_1, \ldots, \tilde{x}_n)$.

Dăm în continuare încă cîteva proprietăți importante ale integralei Lebesgue.

Pentru orice $A \in \mathfrak{K}$ definim integrala lui $\tilde{x}$ pe A în raport cu P, pe care o notăm prin $\int_A \tilde{x}(\omega)\, dP(\omega)$, ca fiind integrala variabilei aleatoare $\tilde{x} \cdot \chi_A$. Avem deci prin definiție

$$\int_A \tilde{x}(\omega)\, dP(\omega) = \int_E \tilde{x}(\omega)\, \chi_A(\omega)\, dP(\omega). \quad (5.1.13)$$

Este clar că pentru orice $A \in \mathfrak{K}$ putem scrie

$$P(A) = \int_E \chi_A(\omega)\, dP(\omega) = \int_A dP(\omega). \quad (5.1.14)$$

Dacă ținem seama și de teorema 5.1.1, atunci pentru orice $B \in \mathfrak{B}$, avem

$$P(\tilde{x} \in B) = P(\tilde{x}^{-1}(B)) = \int_{\tilde{x}^{-1}(B)} dP(\omega) = \int_E \chi_{\tilde{x}^{-1}(B)}(\omega)\, dP(\omega) =$$

$$= \int_E \chi_B(\tilde{x}(\omega))\, dP(\omega) = \int_R \chi_B(x)\, dF(x)$$

și deci

$$P(\tilde{x} \in B) = \int_B dF(x), \quad (5.1.15)$$

formulă utilă pentru calculul probabilității ca $\tilde{x}$ să aparțină unei mulțimi $B \in \mathfrak{B}$.

Scanned with OKEN Scanner

Procedînd analog și utilizînd rezultatul (5.1.12), obținem formula utilă

$$P((\tilde{x}_1, \ldots, \tilde{x}_n) \in B) = \int_B dF(x_1, \ldots, x_n), \tag{5.1.16}$$

pentru orice $B \in \mathfrak{B}^n$. Această formulă ne permite calculul probabilității ca vectorul aleator $(\tilde{x}_1, \ldots, \tilde{x}_n)$ să aparțină unei mulțimi $B \in \mathfrak{B}^n$, cu ajutorul funcției de repartiție $F(x_1, \ldots, x_n)$ a acestui vector aleator.

Proprietățile integralei Lebesgue care se utilizează mai frecvent sînt următoarele:

(*a*) (liniaritate). *Dacă* $\tilde{x}$ *și* $\tilde{y} \in \mathfrak{L}(E, \mathcal{K}, P)$, *iar* a *și* b *sînt numere reale arbitrare, atunci*

$$\int_A (a\tilde{x} + b\tilde{y})\, dP = a\int_A \tilde{x}\, dP + b\int_A \tilde{y}\, dP \tag{5.1.17}$$

pentru orice $A \in \mathcal{K}$;

(*b*) (aditivitate). *Dacă* $(A_i)_{i \in I} \subset \mathcal{K}$ *este cel mult numărabilă și* $A_i \cap A_j = \varnothing$ *pentru* $i \neq j$, *atunci*

$$\int_{\bigcup_{i \in I} A_i} \tilde{x}\, dP = \sum_{i \in I} \int_{A_i} \tilde{x}\, dP; \tag{5.1.18}$$

(*c*) (monotonie). *Dacă* $\tilde{x} \leqslant \tilde{y}$ *aproape sigur pe* $A \in \mathcal{K}$, *atunci*

$$\int_A \tilde{x}\, dP \leqslant \int_A \tilde{y}\, dP; \tag{5.1.19}$$

(*d*) (medie). *Dacă* $a \leqslant \tilde{x} \leqslant b$ *aproape sigur pe* $A \in \mathcal{K}$, *atunci*

$$aP(A) \leqslant \int_A \tilde{x}\, dP \leqslant bP(A); \tag{5.1.20}$$

(*e*) (inegalitatea modulului). *Pentru orice* $A \in \mathcal{K}$ *avem*

$$\left| \int_A \tilde{x}\, dP \right| \leqslant \int_A |\tilde{x}|\, dP; \tag{5.1.21}$$

Scanned with OKEN Scanner

(*f*) (lema lui Fatou). *Dacă* $(\widetilde{x}_n)$ *este un șir de variabile aleatoare pentru care* $\widetilde{x}_n \geqslant 0$ *aproape sigur pe* A, *atunci*

$$\int_A \left(\lim_n \widetilde{x}_n\right) dP \leqslant \lim_n \int_A \widetilde{x}_n \, dP. \tag{5.1.22}$$

Vom încheia prezentarea succintă a integralei Lebesgue cu încă cîteva formule utile privind funcțiile de repartiție.

Să considerăm vectorul aleator $(\widetilde{x}_1, ..., \widetilde{x}_n)$ pe cîmpul borelian de probabilitate $(E, \mathcal{K}, P)$ și fie $g_i : R^n \to R$ funcții măsurabile, $1 \leqslant i \leqslant m$. Presupunînd cunoscută funcția de repartiție $F_{\widetilde{\mathbf{x}}}$ a vectorului aleator $\widetilde{\mathbf{x}} = (\widetilde{x}_1, ..., \widetilde{x}_n)$, ne punem problema stabilirii unei formule pentru calculul funcției de repartiție $F_{\widetilde{z}}$ a vectorului aleator $\widetilde{\mathbf{z}} = (\widetilde{z}_1, ..., \widetilde{z}_m)$, unde $\widetilde{z}_i(\omega) = g_i(\widetilde{\mathbf{x}}(\omega))$, $1 \leqslant i \leqslant m$. Pentru aceasta, să observăm mai întîi că

$$F_{\widetilde{z}}(z_1, ..., z_m) = P(\widetilde{z}_1 < z_1, ..., \widetilde{z}_m < z_m) = P(\widetilde{\mathbf{x}} \in D),$$

unde

$$D = \{(x_1, ..., x_n) : g_i(x_1, ..., x_n) < z_i, \ 1 \leqslant i \leqslant m\}. \tag{5.1.23}$$

Utilizînd formula (5.1.16), rezultă că putem scrie

$$F_{\widetilde{z}}(z_1, ..., z_m) = \int_D dF_{\widetilde{\mathbf{x}}}(x_1, ..., x_n), \tag{5.1.24}$$

care este formula căutată.

În cazul particular cînd $F_{\widetilde{\mathbf{x}}}$ are o densitate de probabilitate $f_{\widetilde{\mathbf{x}}}$, relația precedentă devine

$$F_{\widetilde{z}}(z_1, ..., z_m) = \int_D f_{\widetilde{\mathbf{x}}}(x_1, ..., x_n) \, dx_1 ... \, dx_n. \tag{5.1.25}$$

Să presupunem în plus că $m = n$ și că transformările $y_i = g_i(x_1, ..., x_n)$, $1 \leqslant i \leqslant n$, ale lui R^n în el însuși sînt inversabile, iar transformările inverse $x_j = h_j(y_1, ..., y_n)$, $1 \leqslant j \leqslant n$, sînt funcții continue. Presupunem de asemenea că derivatele

parțiale $\frac{\partial h_j}{\partial x_i}$, $1 \leqslant i, j \leqslant n$ există și sînt continue, iar jacobianul

$$J = \frac{D(h_1, \ldots, h_n)}{D(y_1, \ldots, y_n)} = \left(\frac{\partial h_i}{\partial y_j}\right) 1 \leqslant i, \quad j \leqslant n$$

este nenul. În aceste condiții, formula (5.1.25) devine

$$F_{\tilde{z}}(z_1, \ldots, z_n) = \int_{-\infty}^{z_1} \ldots \int_{-\infty}^{z_n} f_{\tilde{x}}(h_1(\mathrm{y}), \ldots, h_n(\mathrm{y})) \, |J| \, \mathrm{d}y_1 \ldots \mathrm{d}y_n.$$

Formula obținută mai sus arată că $\tilde{\mathbf{z}}$ are o densitate de probabilitate $f_{\tilde{z}}$, care este dată de relația

$$f_{\tilde{z}}(z_1, \ldots, z_n) = f_{\tilde{x}}(h_1(\mathrm{z}), \ldots, h_n(\mathrm{z})) \, |J|. \qquad (5.1.26)$$

Vom prezenta acum o formulă pentru calculul funcției de repartiție a unei sume de variabile aleatoare independente. Pentru simplitate vom începe cu cazul a două variabile aleatoare independente $\tilde{x}_1$ și $\tilde{x}_2$, ale căror funcții de repartiție sînt F_1 și F_2. Deoarece $\tilde{x}_1$ și $\tilde{x}_2$ sînt variabile aleatoare independente, rezultă că funcția de repartiție $F(x_1, x_2)$ a vectorului aleator $(\tilde{x}_1, \tilde{x}_2)$ se scrie sub forma

$$F(x_1, x_2) = F_1(x_1) \, F_2(x_2).$$

Funcția de repartiție a variabilei aleatoare $\tilde{x}_1 + \tilde{x}_2$ o vom nota prin $F_{(2)}(x)$. Este clar că

$$F_{(2)}(x) = P(\tilde{x}_1 + \tilde{x}_2 < x) = \int_{D(x)} \mathrm{d}F(x_1, x_2),$$

unde

$$D(x) = \{(x_1, x_2) : x_1 + x_2 < x\},$$

conform formulei (5.1.24). Din proprietățile integralei Stieltjes rezultă*

$$F_{(2)}(x) = \int_{-\infty}^{+\infty} F_1(x - y) \, \mathrm{d}F_2(y) = \int_{-\infty}^{+\infty} F_2(x - y) \, \mathrm{d}F_1(y). \qquad (5.1.27)$$

* Vezi § 8.4.

Valoarea comună a celor două integrale din (5.1.27) este numită *produsul de compoziție* al funcțiilor de repartiție F_1 și F_2 și este notată prin $(F_1 * F_2)(x)$. Este clar că produsul de compoziție este comutativ și asociativ (vezi și § 8.4). Deci se poate defini produsul de compoziție $F_1 * \ldots * F_n$ a n funcții de repartiție ale variabilelor aleatoare $\tilde{x}_1, \ldots, \tilde{x}_n$ pe $(E, \mathcal{K}, P)$.

Din cele de mai sus rezultă că funcția de repartiție $F_{(n)}$ a sumei $\tilde{x}_1 + \ldots + \tilde{x}_n$ a n variabile aleatoare independente este produsul de compoziție al funcțiilor lor de repartiție, adică

$$F_{(n)} = F_1 * \ldots * F_n. \tag{5.1.28}$$

Evident, dacă funcțiile de repartiție F_1 și F_2 au densități de probabilitate pe f_1 și f_2, atunci formula (5.1.27) devine

$$F_{(2)}(x) = \int_{-\infty}^{+\infty} F_1(x-y)\, f_2(y)\, \mathrm{d}y,$$

ceea ce arată că $F_{(2)}$ are de asemenea o densitate de probabilitate $f_{(2)}$ dată de formula

$$f_{(2)}(x) = \int_{-\infty}^{+\infty} f_1(x-y)\, f_2(y)\, \mathrm{d}y = \int_{-\infty}^{+\infty} f_2(x-y) f_1(y)\, \mathrm{d}y. \tag{5.1.29}$$

Valoarea comună a celor două integrale din (5.1.29) este numită *produsul de compoziție* al densităților de probabilitate f_1 și f_2 și este notată prin $f_1 * f_2$. Rezultă atunci, ca mai sus, că densitatea de probabilitate $f_{(n)}$ a sumei $\tilde{x}_1 + \ldots + \tilde{x}_n$ a n variabile aleatoare independente avînd densitățile de probabilitate $f_1, \ldots, f_n$ este produsul de compoziție al acestor densități, adică

$$f_{(n)} = f_1 * \ldots * f_n. \tag{5.1.30}$$

5.2. MEDIA UNEI VARIABILE ALEATOARE

Să considerăm un cîmp borelian de probabilitate $(E, \mathcal{K}, P)$ și o variabilă aleatoare integrabilă $\tilde{x}$ pe acest cîmp.

Scanned with OKEN Scanner

Definiția 5.2.1. *Integrala Lebesgue a variabilei aleatoare* $\tilde{x}$ *este numită media sa și este notată prin* $M(\tilde{x})$.

Este clar că

$$M(\tilde{x}) = \int_E \tilde{x}(\omega)\, dP(\omega) = \int_R x\, dF(x), \tag{5.2.1}$$

unde F este funcția de repartiție a lui $\tilde{x}$.

Dacă variabila aleatoare $\tilde{x}$ este discretă cu mulțimea valorilor $(x_i)_{i \in I}$, atunci media sa este

$$M(\tilde{x}) = \sum_{i \in I} x_i f(x_i), \tag{5.2.2}$$

unde f este funcția de masă a lui $\tilde{x}$. Cu alte cuvinte, media unei variabile aleatoare discrete este media ponderată a valorilor pe care le ia această variabilă, ponderile fiind chiar probabilitățile cu care aceste valori sînt luate. Această observație justifică denumirea dată numărului $M(\tilde{x})$ definit de formula (5.2.1). Uneori, media lui $\tilde{x}$ este numită și *valoarea medie* a lui $\tilde{x}$.

Dacă $\tilde{x}$ este o variabilă aleatoare continuă, avînd densitatea de probabilitate f, atunci media sa se poate scrie sub forma

$$M(\tilde{x}) = \int_{-\infty}^{+\infty} x\, f(x)\, dx. \tag{5.2.3}$$

O interpretare geometrică interesantă a valorii medii se obține din rezultatul următor.

Teorema 5.2.1. *Dacă* $\tilde{x}$ *este o variabilă aleatoare integrabilă pe* $(E, \mathfrak{K}, P)$, *iar* F *este funcția sa de repartiție, atunci*

$$\lim_{x \to +\infty} x(1 - F(x)) = 0, \tag{5.2.4}$$

$$\lim_{x \to -\infty} xF(x) = 0 \tag{5.2.5}$$

și

$$M(\tilde{x}) = -\int_{-\infty}^{0} F(x)\, dx + \int_{0}^{\infty} (1 - F(x))\, dx. \tag{5.2.6}$$

Demonstrație. Să arătăm mai întîi că au loc rezultatele (5.2.4) și (5.2.5). Deoarece $\widetilde{x}$ este integrabilă, rezultă că și $|\widetilde{x}|$ este integrabilă și deci

$$\lim_{x\to+\infty}\int_x^{\infty}|t|\,\mathrm{d}F(t)=0 \tag{5.2.7}$$

și

$$\lim_{x\to-\infty}\int_{-\infty}^{x}|t|\,\mathrm{d}F(t)=0. \tag{5.2.8}$$

Pentru $x>0$ avem

$$x(1-\mathrm{F}(x))=x\int_x^{\infty}\mathrm{d}F(t)\leqslant\int_x^{\infty}t\,\mathrm{d}F\,(t), \tag{5.2.9}$$

iar pentru $x<0$ avem

$$|x|F(x)=|x|\int_{-\infty}^{x}\mathrm{d}F(x)\leqslant\int_{-\infty}^{x}|t|\,\mathrm{d}F(t). \tag{5.2.10}$$

Trecînd la limită în (5.2.9) pentru $x\to+\infty$ și ținînd seama de (5.2.7), se obține relația (5.2.4).

Analog, trecînd la limită în (5.2.10) pentru $x\to-\infty$ și ținînd seama de (5.2.8) se obține (5.2.5).

Pentru a demonstra relația (5.2.6), să observăm mai întîi că, printr-o integrare prin părți, obținem

$$\int_0^x t\,\mathrm{d}F(t)+\int_0^x F(t)\,\mathrm{d}t=tF(t)\Big|_0^x=xF(x)$$

sau

$$\int_0^x t\,\mathrm{d}F(t)=-x(1-F(x))+\int_0^x(1-F(t))\,\mathrm{d}t. \tag{5.2.11}$$

Scanned with OKEN Scanner

Făcînd o nouă integrare prin părți obținem

$$\int_x^0 t\,\mathrm{d}F(t) + \int_x^0 F(t)\,\mathrm{d}t = tF(t)\Big|_x^0 = -xF(x)$$

sau

$$\int_x^0 t\,\mathrm{d}F(t) = -xF(x) - \int_x^0 F(t)\,\mathrm{d}t. \tag{5.2.12}$$

Trecînd la limită în relația (5.2.11) pentru $x \to +\infty$ și ținînd seama de relația (5.2.4) obținem

$$\int_0^\infty x\,\mathrm{d}F(x) = \int_0^\infty (1 - F(x))\,\mathrm{d}x. \tag{5.2.13}$$

Analog, ținînd seama de relația (5.2.5), din relația (5.2.12) obținem

$$\int_{-\infty}^0 x\,\mathrm{d}F(x) = -\int_{-\infty}^0 F(x)\,\mathrm{d}x \tag{5.2.14}$$

prin trecere la limită pentru $x \to -\infty$.

Din relațiile (5.1.13) și (5.1.14) se deduce ușor rezultatul căutat.

Observație. Din formula (5.2.6) rezultă că media variabilei aleatoare $\tilde{x}$ este suma ariilor haşurate din fig. 5.2.1, aria din stînga axei verticale fiind luată cu semn schimbat, iar cea din dreapta cu semnul plus. Această observație este utilă în unele cazuri pentru a ocoli calcule complicate în determinarea mediei. În particular, dacă cele două arii haşurate sînt egale, media este nulă.

Proprietățile mediei unei variabile aleatoare integrabile se deduc ușor din proprietățile integralei Lebesgue sau ale integralei Stieltjes.

Scanned with OKEN Scanner

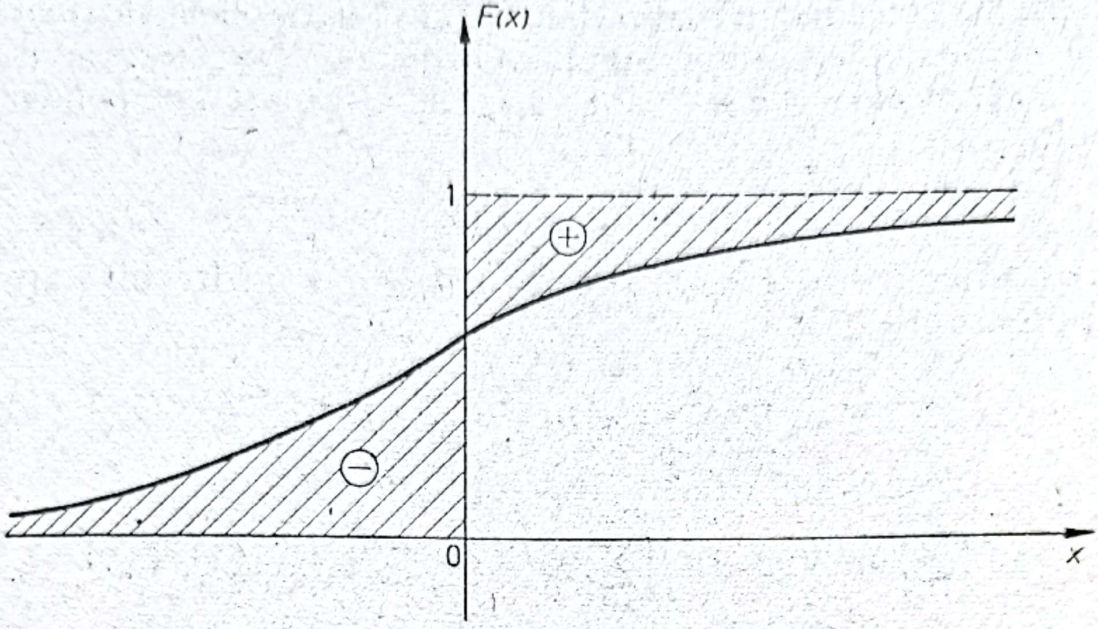


Fig. 5.2.1

Teorema 5.2.2. *Media unei variabile aleatoare are următoarele proprietăți:*

(a) Dacă $\tilde{x}$ *este constantă aproape sigur (adică* $\tilde{x} = a$ *aproape sigur), atunci media lui* $\tilde{x}$ *este acea constantă, adică* $M(\tilde{x}) = a$;

(b) Dacă $\tilde{x}, \tilde{y} \in \mathfrak{L}(E, \mathfrak{K}, P)$ *și dacă* $\tilde{x} \leqslant \tilde{y}$ *aproape sigur, atunci avem* $M(\tilde{x}) \leqslant M(\tilde{y})$;

(c) Dacă $\tilde{x}, \tilde{y} \in \mathfrak{L}(E, \mathfrak{K}, P)$ *și* $a, b \in R$, *atunci* $a\tilde{x} + b\tilde{y} \in \mathfrak{L}$ și

$$M(a\tilde{x} + b\tilde{y}) = aM(\tilde{x}) + bM(\tilde{y}); \tag{5.2.15}$$

(d) Dacă $\tilde{x}$ și $\tilde{y}$ *sînt variabile aleatoare integrabile și independente, atunci* $\tilde{x}\tilde{y}$ *este integrabilă și*

$$M(\tilde{x}\tilde{y}) = M(\tilde{x})\, M(\tilde{y}); \tag{5.2.16}$$

(e) Dacă $\tilde{x}$ *este o variabilă aleatoare integrabilă, iar* $g: R \to R$ *este convexă, atunci*

$$g(M(\tilde{x})) \leqslant M(g(\tilde{x})). \tag{5.2.17}$$

Demonstrație. (*a*), (*b*) și (*c*) sînt consecințe imediate ale proprietăților integralei Lebesgue.

(*d*). Să considerăm mai întîi cazul cînd variabilele aleatoare $\tilde{x}$ și $\tilde{y}$ sînt simple, prima luînd valorile $x_1, \ldots, x_m$, iar a doua $y_1, \ldots, y_n$. Deoarece $\tilde{x}$ și $\tilde{y}$ sînt variabile aleatoare independente, rezultă că

$$P(\tilde{x} = x_i,\ \tilde{y} = y_j) = P(\tilde{x} = x_i)\ P(\tilde{y} = y_i),$$

pentru orice i, $1 \leqslant i \leqslant m$, și orice j, $1 \leqslant j \leqslant n$. Rezultă atunci că putem scrie

$$M(\tilde{x}\tilde{y}) = \sum_{i=1}^{m}\sum_{j=1}^{n} x_i y_j P(\tilde{x} = x_i)\ P(\tilde{y} = y_j) =$$

$$= \left(\sum_{i=1}^{m} x_i P(\tilde{x} = x_i)\right)\left(\sum_{j=1}^{n} y_j P(\tilde{y} = y_j)\right) = M(\tilde{x})\, M(\tilde{y}),$$

ceea ce demonstrează afirmația în acest caz.

Să considerăm acum cazul cînd variabilele aleatoare independente $\tilde{x}$ și $\tilde{y}$ sînt nenegative. Conform unui rezultat cunoscut (vezi teorema 4.1.4), rezultă că există un șir crescător de variabile aleatoare simple $(\tilde{x}_n)$ convergent spre $\tilde{x}$ și un șir crescător de variabile aleatoare simple $(\tilde{y}_n)$ convergent spre $\tilde{y}$. Din independența variabilelor aleatoare $\tilde{x}$ și $\tilde{y}$ rezultă că evenimentele

$$A_n^i = \{\omega : (i-1)/2^n \leqslant \tilde{x}(\omega) < i/2^n\}$$

și

$$B_n^j = \{\omega : (j-1)/2^n \leqslant \tilde{y}(\omega) < j/2^n\}$$

sînt independente, ceea ce implică independența variabilelor aleatoare simple $\tilde{x}_n$ și $\tilde{y}_n$. Din prima parte a demonstrației rezultă că

$$M(\tilde{x}_n\tilde{y}_n) = M(\tilde{x}_n)\, M(\tilde{y}_n).$$

Deoarece șirul $(\tilde{x}_n\tilde{y}_n)$ este crescător și converge spre $\tilde{x}\tilde{y}$, rezultă că $\tilde{x}\tilde{y}$ este integrabilă și, ținînd seama și de proprietatea de convergență monotonă a integralei Lebesgue, obținem

$$M(\tilde{x}\tilde{y}) = M(\tilde{x})\, M(\tilde{y}).$$

Am demonstrat deci afirmația (*d*) pentru variabile aleatoare independente nenegative.

Scanned with OKEN Scanner

Fie $\tilde{x}$ și $\tilde{y}$ variabile aleatoare independente și integrabile. Deoarece se poate scrie $\tilde{x} = \tilde{x}^+ - \tilde{x}^-$ și $\tilde{y} = \tilde{y}^+ - \tilde{y}^-$, rezultă că variabilele

$$\tilde{x}' = \tilde{x}^+ \text{ sau } \tilde{x}^- \text{ sau } |\tilde{x}|$$

și

$$\tilde{y}' = \tilde{y}^+ \text{ sau } \tilde{y}^- \text{ sau } |\tilde{y}|$$

sînt de asemenea independente și nenegative. Rezultă integrabilitatea variabilei $\tilde{x}\tilde{y}$ și în plus

$$M(\tilde{x}\tilde{y}) = M(\tilde{x}^+ - \tilde{x}^-)(\tilde{y}^+ - \tilde{y}^-)) = M(\tilde{x}^+\tilde{y}^+) -$$
$$- M(\tilde{x}^+\tilde{y}^-) - M(\tilde{x}^-\tilde{y}^+) + M(\tilde{x}^-\tilde{y}^-) = M(\tilde{x}^+)\,M(\tilde{y}^+) -$$
$$- M(\tilde{x}^+)\,M(\tilde{y}^-) - M(\tilde{x}^-)\,M(\tilde{y}^+) + M(\tilde{x}^-)\,M(\tilde{y}^-) =$$
$$= (M(\tilde{x}^+) - M(\tilde{x}^-))\,(M(\tilde{y}^+) - M(\tilde{y}^-)) = M(\tilde{x})\,M(\tilde{y})$$

și deci proprietatea (d) este demonstrată în cazul general.

(e). Dacă g este o funcție convexă definită pe R, atunci pentru orice $x_0 \in R$ există un număr real u_0 astfel încît

$$g(x) - g(x_0) \geqslant (x - x_0)\,u_0$$

pentru orice $x \in R$. Punînd $x_0 = M(\tilde{x})$ și înlocuind x cu $\tilde{x}$ obținem

$$g(\tilde{x}) - g(M(\tilde{x})) \geqslant (\tilde{x} - M(\tilde{x}))\,u_0.$$

Luînd valoarea medie în această inegalitate și ținînd seama de proprietățile (a), (b) și (c) obținem

$$M(g(\tilde{x})) - g(M(\tilde{x})) \geqslant u_0(M(\tilde{x}) - M(\tilde{x})) = 0,$$

de unde se obține rezultatul dorit.

Observații. Este clar că proprietatea (c) se generalizează în modul următor. Dacă $\tilde{x}_i \in L(E, \mathfrak{K}, P)$ pentru $1 \leqslant i \leqslant n$, și $a_i \in R$, $1 \leqslant i \leqslant n$, atunci

$$M\left(\sum_{i=1}^{n} a_i\tilde{x}_i\right) = \sum_{i=1}^{n} a_i M(\tilde{x}_i).$$

Aceeași observație este valabilă în legătură cu proprietatea (*d*). Mai exact, dacă $\tilde{x}_i \in L(E, \mathfrak{K}, P)$ pentru $1 \leqslant i \leqslant n$, și dacă $\tilde{x}_1, \ldots, \tilde{x}_n$ sînt independente, atunci

$$M\left(\prod_{i=1}^{n} \tilde{x}_i\right) = \prod_{i=1}^{n} M(\tilde{x}_i).$$

Se poate de asemenea vedea că în afirmația (*d*) se poate înlocui ipoteza $\tilde{x}, \tilde{y} \in L(E, \mathfrak{K}, P)$ prin ipoteza $\tilde{x}\tilde{y} \in L(E, \mathfrak{K}, P)$, obținînd de aici integrabilitatea variabilelor aleatoare $\tilde{x}$ și $\tilde{y}$, în cazul cînd nici una dintre ele nu este nulă aproape sigur.

5.3. DISPERSIA UNEI VARIABILE ALEATOARE

Fie $\tilde{x}$ o variabilă aleatoare pe cîmpul borelian de probabilitate $(E, \mathfrak{K}, P)$. Dacă $\tilde{x}$ este integrabilă, atunci putem considera variabila aleatoare

$$\tilde{y} = \tilde{x} - M(\tilde{x}),$$

numită și *variabilă aleatoare abatere;* această variabilă aleatoare descrie împrăștierea valorilor luate de variabila aleatoare $\tilde{x}$ în jurul valorii sale medii.

Definiția 5.3.1. *Dacă variabila aleatoare* $\tilde{y}^2 \in L(E, \mathfrak{K}, P)$, *atunci media variabilei aleatoare* $\tilde{y}^2$ *este numită dispersia sau varianța variabilei aleatoare* $\tilde{x}$ *și este notată prin* $V(\tilde{x})$.

Pentru notarea dispersiei se mai folosesc și simbolurile $D^2(\tilde{x})$, $\sigma^2(\tilde{x})$ sau σ^2, atunci cînd nu există pericol de confuzie. Numărul $\sqrt{V(\tilde{x})} = \sqrt{D^2(\tilde{x})} = \sqrt{\sigma^2(\tilde{x})}$, care se mai notează uneori prin $D(\tilde{x})$, $\sigma(\tilde{x})$ sau σ, atunci cînd nu există pericol de confuzie, este numit *abaterea medie pătratică* sau *abaterea standard* a variabilei aleatoare $\tilde{x}$.

Varianța (dispersia) și abaterea standard reprezintă măsuri ale împrăștierii valorilor variabilei aleatoare în jurul valorii sale medii. Dacă $V(\tilde{x}) \neq 0$, vom vorbi uneori de *precizia* variabilei aleatoare $\tilde{x}$, care este definită ca inversa dispersiei; precizia variabilei aleatoare $\tilde{x}$ este deci $1/V(\tilde{x})$. Deoarece dispersia este

Scanned with OKEN Scanner

o măsură a incertitudinii asupra valorilor pe care le ia o variabilă aleatoare, rezultă că precizia poate fi interpretată ca o măsură asupra *certitudinii* sau *informației* asupra lui $\tilde{x}$.

Din definiția dispersiei rezultă că

$$V(\tilde{x}) = M(\tilde{x} - M(\tilde{x}))^2, \tag{5.3.1}$$

de unde obținem ușor formula utilă

$$V(\tilde{x}) = M(\tilde{x}^2) - M^2(\tilde{x}). \tag{5.3.2}$$

Într-adevăr, deoarece

$$V(\tilde{x}) = M(\tilde{x}^2 - 2\tilde{x}M(\tilde{x}) + M^2(\tilde{x})) =$$
$$= M(\tilde{x}^2) - 2M^2(\tilde{x}) + M^2(\tilde{x}),$$

rezultă imediat (5.3.2).

Să observăm de asemenea că avem

$$V(\tilde{x}) = \int_E (\tilde{x} - M(\tilde{x}))^2 \, dP = \int_R (x - M(\tilde{x}))^2 \, dF(x), \tag{5.3.3}$$

ținînd seama de definiția mediei unei variabile aleatoare. În particular, dacă $\tilde{x}$ este discretă cu valorile $(x_i)_{i \in I}$, rezultă

$$V(\tilde{x}) = \sum_{i \in I} (x_i - M(\tilde{x}))^2 f(x_i) = \sum_{i \in I} x_i^2 f(x_i) - M^2(\tilde{x}), \tag{5.3.4}$$

unde f este funcția de masă a lui $\tilde{x}$. Analog, dacă $\tilde{x}$ este continuă, atunci

$$V(\tilde{x}) = \int_R (x - M(\tilde{x}))^2 f(x) \, dx = \int_R x^2 f(x) \, dx - M^2(\tilde{x}), \tag{5.3.5}$$

unde f este densitatea de probabilitate a lui $\tilde{x}$.

Proprietățile dispersiei sînt date în teorema următoare.

Teorema 5.3.1. *(a) Dacă* $\tilde{x} =$ a *aproape sigur, atunci* $V(\tilde{x}) = 0$;

(b) Dacă variabilele aleatoare $\tilde{x}_1, \ldots, \tilde{x}_n$ *sînt două cîte două independente și dacă* $\tilde{x}_i^2 \in L(E, \mathcal{K}, P)$, $1 \leqslant i \leqslant n$, *atunci*

$$V\left(\sum_{i=1}^{n} a_i \tilde{x}_i\right) = \sum_{i=1}^{n} a_i^2 V(\tilde{x}_i), \tag{5.3.6}$$

unde $a_1, \ldots, a_n$ *sînt numere reale.*

Demonstrație. (a). Afirmația rezultă imediat din definiție, ținînd seama că $M(\tilde{x}) = a$ și deci $(\tilde{x} - M(\tilde{x}))^2 = 0$ aproape sigur.

(b). Dacă notăm $m_i = M(\tilde{x}_i)$ și

$$\tilde{x} = \sum_{i=1}^{n} a_i \tilde{x}_i,$$

atunci

$$M(\tilde{x}) = \sum_{i=1}^{n} a_i m_i.$$

Prin urmare putem scrie

$$(\tilde{x} - M(x))^2 = \left(\sum_{i=1}^{n} a_i (\tilde{x}_i - m_i)\right)^2 =$$

$$= \sum_{i=1}^{n} a_i^2 (\tilde{x}_i - m_i)^2 + 2 \sum_{1 \leqslant i < j \leqslant n} a_i a_j (\tilde{x}_i - m_i)(\tilde{x}_j - m_j)$$

și deci

$$M((\tilde{x} - M(\tilde{x}))^2) = \sum_{i=1}^{n} a_i^2 M((\tilde{x}_i - m_i)^2) +$$

$$+ 2 \sum_{1 \leqslant i < j \leqslant n} a_i a_j M((\tilde{x}_i - m_i)(\tilde{x}_j - m_j)), \tag{5.3.7}$$

ținînd seama de proprietățile mediei.

Să observăm însă că $\tilde{x}_i$ și $\tilde{x}_j$ sînt independente pentru $i \neq j$ și deci aceeași proprietate au și variabilele aleatoare $\tilde{x}_i - m_i$ și $\tilde{x}_j - m_j$. Ținînd din nou seama de proprietățile mediei rezultă că

$$M((\tilde{x}_i - m_i)(\tilde{x}_j - m_j)) = (M(\tilde{x}_i - m_i))(M(\tilde{x}_j - m_j)) = 0. \tag{5.3.8}$$

Din (5.3.7) și (5.3.8) se deduce ușor rezultatul căutat.

Scanned with OKEN Scanner

5.4. MOMENTE. MOMENTE CENTRATE

Să considerăm un cîmp borelian de probabilitate $(E, \mathcal{K}, P)$ și o variabilă aleatoare $\tilde{x}$ pe acest cîmp.

Definiția 5.4.1. *Dacă pentru* $p \in N$, *variabila aleatoare* $\tilde{x}^p$ *este integrabilă, atunci integrala sa este numită momentul de ordinul* p *al variabilei aleatoare* $\tilde{x}$ *și este notat prin* $M_p(\tilde{x})$.

Din această definiție rezultă că putem scrie

$$M_p(\tilde{x}) = \int_E \tilde{x}^p \, dP(\omega) = \int_R x^p \, dF(x), \tag{5.4.1}$$

unde F este funcția de repartiție a variabilei aleatoare $\tilde{x}$.

Dacă pentru $r \in R$, $r > 0$, variabila aleatoare $|\tilde{x}|^r$ este integrabilă, atunci integrala sa este numită *momentul absolut de ordinul* r al variabilei aleatoare $\tilde{x}$ și este notat prin $\bar{M}_r(\tilde{x})$. Rezultă că putem scrie

$$\bar{M}_r(\tilde{x}) = \int_E |\tilde{x}(\omega)|^r \, dP(\omega) = \int_R |x|^r \, dF(x). \tag{5.4.2}$$

Din definiție rezultă ușor că

$$\bar{M}_r(\tilde{x}) = M_r(|\tilde{x}|). \tag{5.4.3}$$

Vom nota prin $\mathcal{L}^r(E, \mathcal{K}, P)$ mulțimea variabilelor aleatoare pe cîmpul de probabilitate $(E, \mathcal{K}, P)$ cu proprietatea că $\tilde{x}^r$ este integrabilă pe E în raport cu P.

Dacă $\tilde{x}$ este o variabilă aleatoare discretă cu mulțimea valorilor $(x_i)_{i \in I}$, atunci momentul de ordinul p al variabilei aleatoare $\tilde{x}$ este

$$M_p(x) = \sum_{i \in I} x_i^p f(x_i),$$

unde f este funcția de masă a lui $\tilde{x}$.

Analog, dacă $\tilde{x}$ este o variabilă aleatoare continuă, atunci momentul de ordinul p al acestei variabile aleatoare este

$$M_p(\tilde{x}) = \int_R x^p f(x) \, dx,$$

unde f este densitatea de probabilitate a lui $\tilde{x}$.

Scanned with OKEN Scanner

Este clar că momentul de ordinul întîi coincide cu media variabilei aleatoare, adică $M_1(\tilde{x}) = M(\tilde{x})$.

Definiția 5.4.2. *Vom numi moment centrat de ordinul* p *al variabilei aleatoare* $\tilde{x}$ *și îl vom nota prin* $V_p(\tilde{x})$, *momentul de ordinul* p *al variabilei aleatoare abatere* $\tilde{y} = \tilde{x} - M(\tilde{x})$.

Avem deci prin definiție

$$V_p(\tilde{x}) = M_p(\tilde{x} - M(\tilde{x})). \tag{5.4.4}$$

Este clar că $V_0(\tilde{x}) = 1$ și $V_1(\tilde{x}) = 0$ pentru orice variabilă aleatoare $\tilde{x}$. Cu alte cuvinte, primul moment centrat semnificativ este $V_2(\tilde{x})$, care coincide evident cu dispersia $V(\tilde{x})$.

Între momentele și momentele centrate ale unei variabile aleatoare se pot stabili ușor relații utile. Astfel, deoarece

$$(\tilde{x} - M(\tilde{x}))^p = \sum_{k=0}^{p} (-1)^k C_p^k \tilde{x}^{p-k} M^k(\tilde{x}),$$

prin trecere la valoarea medie rezultă că

$$V_p(\tilde{x}) = \sum_{k=0}^{p} (-1)^k C_p^k M_{p-k}(\tilde{x})\, M^k(\tilde{x}). \tag{5.4.5}$$

Analog, trecînd la valoarea medie în relația evidentă

$$\tilde{x}^p = ((\tilde{x} - M(\tilde{x})) + M(\tilde{x}))^p = \sum_{k=0}^{p} C_p^k (\tilde{x} - M(\tilde{x}))^{p-k} M^k(\tilde{x}),$$

obținem

$$M_p(\tilde{x}) = \sum_{k=0}^{p} C_p^k V_{p-k}(\tilde{x})\, M^k(\tilde{x}). \tag{5.4.6}$$

Proprietățile cele mai importante ale momentelor sînt date în teorema următoare.

Teorema 5.4.1. *(a) Dacă* $\tilde{x} = a$ *aproape sigur, atunci* $M_p(\tilde{x}) = a^p$ *pentru orice* $p \in N$;

(b) Dacă $\overline{M}_r(\tilde{x}) < +\infty$, *atunci* $\overline{M}_{r'}(\tilde{x}) < +\infty$ *pentru orice* $r' \leqslant r$ *și* $M_p(\tilde{x}) < +\infty$ *pentru orice* $p \in N$ *cu* $p \leqslant r$;

(c) Dacă $\tilde{x}, \tilde{y} \in \mathfrak{L}^r(E, \mathcal{K}, P)$, *atunci* $\tilde{x} + \tilde{y} \in \mathfrak{L}^r(E, \mathcal{K}, P)$ *și*

$$\overline{M}_r(\tilde{x} + \tilde{y}) \leqslant c_r \overline{M}_r(\tilde{x}) + c_r \overline{M}_r(\tilde{y}), \tag{5.4.7}$$

unde c_r *este egal cu* 1 *sau cu* 2^{r-1}, *după cum* $r \leqslant 1$ *sau* $r > 1$;

(d) (Inegalitatea lui Hölder) Dacă $\tilde{x} \in \mathfrak{L}^r(E, \mathcal{K}, P)$ *și* $\tilde{y} \in \mathfrak{L}^r(E, \mathcal{K}, P)$, *unde* $r > 1$ *și* $1/r + 1/s = 1$, *atunci* $\tilde{x}\tilde{y} \in \mathfrak{L}(E, \mathcal{K}, P)$ *și*

$$\overline{M}(\tilde{x}\tilde{y}) \leqslant (\overline{M}_r(\tilde{x}))^{1/r} (\overline{M}_s(\tilde{y}))^{1/s}; \tag{5.4.8}$$

(e) (Inegalitatea lui Minkowsky). Dacă $\tilde{x}, \tilde{y} \in \mathfrak{L}^r(E, \mathcal{K}, P)$ *și* $r \geqslant 1$, *atunci* $\tilde{x} + \tilde{y} \in \mathfrak{L}^r(E, \mathcal{K}, P)$ *și*

$$(\overline{M}_r(\tilde{x} + \tilde{y}))^{1/r} \leqslant (\overline{M}_r(\tilde{x}))^{1/r} + (\overline{M}_r(\tilde{y}))^{1/r}; \tag{5.4.9}$$

(f) Dacă $\tilde{x} \in \mathfrak{L}^{r''}(E, \mathcal{K}, P)$ *și* $0 < r' \leqslant r''$, *atunci*

$$(\overline{M}_{r'}(\tilde{x}))^{1/r'} \leqslant (\overline{M}_{r''}(\tilde{x}))^{1/r''}. \tag{5.4.10}$$

Demonstrație. (a) și *(b)* rezultă imediat din proprietățile integralei Lebesgue.

(c) Să considerăm inegalitatea elementară

$$|a + b|^r \leqslant c_r |a|^r + c_r |b|^r,$$

unde a și b sînt numere reale arbitrare, $r > 0$ și c_r este egal cu 1 sau cu 2^{r-1}, după cum $r \leqslant 1$ sau $r > 1$. Înlocuind în această inegalitate pe a prin $\tilde{x}$ și pe b prin $\tilde{y}$ și luînd valoarea medie obținem

$$M(|\tilde{x} + \tilde{y}|^r) \leqslant c_r M(|\tilde{x}|^r) + c_r M(|\tilde{y}|^r),$$

de unde urmează afirmațiile căutate.

(d) Vom porni de la inegalitatea elementară

$$|ab| \leqslant |a|^r/r + |b|^s/s,$$

unde a și b sînt numere reale arbitrare, $r > 1$, iar $1/r + 1/s = 1$.

Pentru demonstrarea inegalității lui Hölder, să observăm mai întîi că, în cazul cînd $\overline{M}_r(\tilde{x}) = 0$ sau $\overline{M}_s(\tilde{y}) = 0$, inegalitatea este evidentă. Presupunînd că $\overline{M}_r(\tilde{x})\, \overline{M}_s(\tilde{y}) \neq 0$ și înlocuind

Scanned with OKEN Scanner

în inegalitatea de mai sus pe a prin $\tilde{x}/(\overline{M}_r(\tilde{x}))^{1/r}$ și pe b prin $\tilde{y}/(\overline{M}_s(\tilde{x}))^{1/s}$ obținem

$$\frac{|\tilde{x}\tilde{y}|}{(\overline{M}_r(\tilde{x}))^{1/r}(\overline{M}_s(\tilde{y})^{1/s}} \leqslant \frac{|\tilde{x}|^r}{r\overline{M}_r(\tilde{x})} + \frac{|\tilde{y}|^r}{s\overline{M}_s(\tilde{y})}.$$

Luînd valoarea medie în această inegalitate obținem ușor rezultatul dorit.

(c) Pentru $r = 1$ inegalitatea este evidentă. Pentru $r > 1$ avem

$$|\tilde{x} + \tilde{y}|^r = |\tilde{x} + \tilde{y}| \cdot |\tilde{x} + \tilde{y}|^{r-1} \leqslant |\tilde{x}| \cdot |\tilde{x} + \tilde{y}|^{r-1} + |\tilde{y}| \cdot |\tilde{x} + \tilde{y}|^{r-1}.$$

Trecînd la valoarea medie obținem

$$\overline{M}_r(\tilde{x} + \tilde{y}) \leqslant M(|\tilde{x}| \cdot |\tilde{x} + \tilde{y}|^{r-1}) + M(|\tilde{y}| \cdot |\tilde{x} + \tilde{y}|^{r-1}).$$

Aplicînd pentru fiecare din cei doi termeni din membrul drept inegalitatea lui Hölder, obținem mai întîi

$$M(|\tilde{x}| \cdot |\tilde{x} + \tilde{y}|^{r-1}) \leqslant (\overline{M}_r(\tilde{x}))^{1/r} (\overline{M}_s((\tilde{x} + \tilde{y})^{r-1}))^{1/s} = (\overline{M}_r(\tilde{x}))^{1/r} (\overline{M}_{s(r-1)} (\tilde{x}+\tilde{y}))^{1/s}.$$

Ținînd seama că $1/s = 1 - 1/r = (r-1)/r$, rezultă că $s(r-1) = r$ și deci

$$(\overline{M}_{s(r-1)}(\tilde{x} + \tilde{y}))^{1/s} = (\overline{M}_r(\tilde{x} + \tilde{y}))^{(r-1)/r}.$$

Avem prin urmare

$$M(|\tilde{x}| \cdot |\tilde{x} + \tilde{y}|^{r-1}) \leqslant (\overline{M}_r(\tilde{x}))^{1/r}(\overline{M}_r(\tilde{x} + \tilde{y}))^{(r-1)/r}.$$

Analog, pentru al doilea termen obținem

$$M(|\tilde{y}| \cdot |\tilde{x} + \tilde{y}|^{r-1}) = (\overline{M}_r(\tilde{y}))^{1/r}(\overline{M}_r(\tilde{x} + \tilde{y}))^{(r-1)/r}$$

și deci

$$\overline{M}_r(\tilde{x} + \tilde{y}) \leqslant \{(\overline{M}_r(\tilde{x}))^{1/r} + (\overline{M}_r(\tilde{y}))^{1/r}\} (\overline{M}_r(\tilde{x} + \tilde{y}))^{(r-1)/r},$$

de unde rezultă imediat inegalitatea căutată.

Scanned with OKEN Scanner

(f) Din proprietatea *(b)* rezultă imediat că $\tilde{y} \in \mathfrak{L}^{r'}(E, \mathcal{K}, P)$. Este clar că pentru $r' = r''$ inegalitatea este imediată. Pentru $r' < r''$, rezultă că $r = r''/r' > 1$. Aplicînd inegalitatea lui Hölder pentru $\tilde{x}^{r'}$ și 1, unde $1/r + 1/s = 1$, rezultă

$$M(|\tilde{x}|^{r'} \cdot 1) \leqslant (\bar{M}_r(\tilde{x}^{r'}))^{1/r} (\bar{M}_s(1))^{1/s}$$

sau

$$\bar{M}_{r'}(\tilde{x}) \leqslant (\bar{M}_{r''}(\tilde{x}))^{r'/r''},$$

de unde rezultă inegalitatea căutată.

Observație. Pentru $r = s = 2$, inegalitatea lui Hölder este cunoscută sub numele de *inegalitatea lui Schwarz:*

$$M(|\tilde{x}\tilde{y}|) \leqslant (\bar{M}_2(\tilde{x})\, \bar{M}_2(\tilde{y}))^{1/2}, \tag{5.4.11}$$

valabilă pentru $\tilde{x}$, $\tilde{y} \in \mathfrak{L}^2(E, \mathcal{K}, P)$.

Vom încheia proprietățile momentelor cu cîteva inegalități utile.

Teorema 5.4.2. *Fie* $\tilde{x}$ *o variabilă aleatoare pe cîmpul borelian de probabilitate* $(E, \mathcal{K}, P)$ *și fie* $g : R \to R$ *o funcție nenegativă, pară, a cărei restricție la* $[0, +\infty)$ *este crescătoare și astfel încît variabila aleatoare* $\tilde{y} = g(\tilde{x})$ *să fie integrabilă. În aceste condiții au loc inegalitățile*:*

$$\frac{M(\tilde{y}) - g(x)}{\text{a.s. sup } \tilde{y}} \leqslant P(|\tilde{x}| \geqslant x) \leqslant \frac{M(\tilde{y})}{g(x)}. \tag{5.4.12}$$

Demonstrație. Fie

$$A = \{\omega : |\tilde{x}(\omega)| \geqslant x\}.$$

Deoarece variabila aleatoare $\tilde{y}$ este integrabilă, avem

$$M(\tilde{y}) = \int_A g(\tilde{x}(\omega))\, dP(\omega) + \int_{A^c} g(\tilde{x}(\omega))\, dP(\omega). \tag{5.4.13}$$

---

* *Supremul aproape sigur* al unei variabile aleatoare $\tilde{y}$ este numărul $M = \text{a.s.sup } \tilde{y}$ definit de condițiile $P(\tilde{y} > M) = 0$ și $P(\tilde{y} \geqslant M - \varepsilon) > 0$ pentru orice $\varepsilon > 0$.

Scanned with OKEN Scanner

Primul termen din membrul drept al relației (5.4.13) verifică dubla inegalitate

$$g(x)\,P(A) \leqslant \int_A g(\tilde{x}(\omega))\,\mathrm{d}P(\omega)) \leqslant (\text{a.s. sup } \tilde{y})\,P(A), \quad (5.4.14)$$

iar al doilea termen verifică dubla inegalitate

$$0 \leqslant \int_{A^c} g(\tilde{x}(\omega))\,\mathrm{d}P(\omega) \leqslant g(x). \quad (5.4.15)$$

Din ultimele trei relații rezultă

$$g(x)\,P(A) \leqslant M(\tilde{y}) \leqslant g(x) + (\text{a.s. sup } \tilde{y})\,P(A), \quad (5.4.16)$$

din care obținem imediat inegalitatea căutată.

Observația 1. Dacă se presupune că g este nenegativă și crescătoare, iar variabila aleatoare $\tilde{y} = g(\tilde{x})$ integrabilă, se poate demonstra ca mai sus dubla inegalitate

$$\frac{M(\tilde{y}) - g(x)}{\text{a.s. sup } \tilde{y}} \leqslant P(\tilde{x} \geqslant x) \leqslant \frac{M(\tilde{y})}{g(x)}, \quad (5.4.17)$$

pentru orice x pentru care $g(x) > 0$.

Observația 2. Dacă se presupune că g este nenegativă, pară, $g(y) \geqslant a > 0$ pentru $y \geqslant x$, iar variabila aleatoare $\tilde{y} = g(\tilde{x})$ este integrabilă, atunci a doua inegalitate din (5.4.12) rămîne adevărată. Într-adevăr, din (5.4.13) se deduce ușor

$$M(\tilde{y}) \geqslant \int_A g(\tilde{x}(\omega))\,\mathrm{d}P(\omega) \geqslant a\,P(A)$$

și deci

$$P(A) \leqslant \frac{M(\tilde{y})}{a}. \quad (5.4.18)$$

Scanned with OKEN Scanner

Observația 3. În cazul particular $g(x) = e^{rx}$ cu $r > 0$, în care g este evident nenegativă și crescătoare, din (5.4.17) obținem

$$\frac{M(e^{r\tilde{x}}) - e^{rx}}{\text{a.s. sup } e^{r\tilde{x}}} \leqslant P(\tilde{x} \geqslant x) \leqslant \frac{M(e^{r\tilde{x}})}{e^{rx}}, \qquad (5.4.19)$$

pentru orice x real. Inegalitatea (5.4.19) este cunoscută sub numele de *inegalitatea lui Kolmogorov.*

Observația 4. În cazul particular $g(x) = |x|^r$ cu $r > 0$, în care g verifică evident condițiile teoremei 5.4.2, obținem

$$\frac{\overline{M}_r(\tilde{x}) - x^r}{\text{a.s. sup}\,|\tilde{x}|^r} \leqslant P(|\tilde{x}| \geqslant x) \leqslant \frac{\overline{M}_r(\tilde{x})}{x^r}, \qquad (5.4.20)$$

pentru $x > 0$, inegalitate cunoscută sub numele de *inegalitatea lui Markov.* Pentru $r = 2$ se obține celebra *inegalitate a lui Cebîșev:*

$$P(|\tilde{x}| \geqslant x) \leqslant \frac{M_2(\tilde{x})}{x^2}, \qquad (5.4.21)$$

valabilă pentru orice $x > 0$.

5.5. MEDIANĂ, CUANTILE, MOD, ASIMETRIE ȘI EXCES

Să considerăm variabila aleatoare $\tilde{x}$ pe cîmpul borelian de probabilitate $(E, \mathcal{K}, P)$ și fie F funcția sa de repartiție. Vom numi *mediană* a lui $\tilde{x}$ și o vom nota prin $m(\tilde{x})$, un număr real cu proprietatea

$$P(\tilde{x} \geqslant m(\tilde{x})) \geqslant 1/2 \leqslant P(\tilde{x} \leqslant m(\tilde{x})) \qquad (5.5.1)$$

sau, echivalent,

$$F(m(\tilde{x})) \leqslant 1/2 \leqslant F(m(\tilde{x}) + 0). \qquad (5.5.2)$$

Scanned with OKEN Scanner

Este clar că orice variabilă aleatoare are cel puțin o mediană. Dacă $\tilde{x}$ este o variabilă aleatoare continuă, atunci rezultă că mediana sa $m(\tilde{x})$ este soluție a ecuației $F(x) = 1/2$. Evident, această ecuație poate avea o soluție unică sau o infinitate de soluții (ultimul caz se întîlnește cînd curba $y = F(x)$ are un segment în comun cu dreapta $y = 1/2$).

Generalizînd noțiunea de mediană ajungem la noțiunea de cuantilă. Pentru orice număr natural $q \geqslant 2$, definim *cuantilele de ordinul* q ca fiind orice numere reale $c_i(\tilde{x})$, $1 \leqslant i \leqslant q$, pentru care

$$P(\tilde{x} \geqslant c_i(\tilde{x})) \geqslant (q-i)/q \text{ și } P(\tilde{x} \leqslant c_i(\tilde{x})) \geqslant i/q \qquad (5.5.3)$$

pentru $1 \leqslant i \leqslant q-1$, sau, echivalent,

$$F(c_i(\tilde{x})) \leqslant i/q \leqslant F(c_i(\tilde{x}) + 0), \quad 1 \leqslant i \leqslant q-1. \qquad (5.5.4)$$

Este clar că, în general, cuantilele nu sînt unic determinate; dacă $\tilde{x}$ este variabilă aleatoare continuă, atunci cuantilele de ordinul q sînt soluții ale ecuațiilor $F(x) = i/q$, $1 \leqslant i \leqslant q$. Este de asemenea clar că cuantila de ordinul doi este chiar mediana, adică $m(\tilde{x}) = c_2(\tilde{x})$.

Fie acum $\tilde{x}$ o variabilă aleatoare continuă și $f(x)$ densitatea sa de probabilitate. Numim *mod* orice punct de maxim local al funcției $f(x)$. Variabilele aleatoare și funcțiile lor de repartiție care au un mod, două moduri, mai multe moduri sînt numite *unimodale, bimodale, multimodale.*

Conceptul de mod se poate defini și pentru variabile aleatoare discrete. Mai exact, dacă $\tilde{x}$ este o variabilă aleatoare discretă cu mulțimea valorilor $(x_i)_{i \in I}$, iar aceste valori sînt în ordine crescătoare, atunci numim *mod* orice valoare x_j cu proprietatea că $f(x_j) > f(x_{j-1})$ și $f(x_j) > f(x_{j+1})$, unde f este funcția de masă a variabilei discrete $\tilde{x}$.

Vom aminti în încheiere și alte două valori numerice, cunoscute sub numele de *asimetrie* și *exces*, care de asemenea pot fi asociate unei variabile aleatoare. Pentru orice variabilă aleatoare $\tilde{x} \in \mathfrak{L}^3(E, \mathfrak{K}, P)$ numim *asimetrie* și o notăm cu γ_1, raportul dintre momentul centrat de ordinul trei și cubul abaterii medii pătratice:

$$\gamma_1 = \frac{V_3(\tilde{x})}{(V_2(\tilde{x}))^{3/2}}. \qquad (5.5.5)$$

Scanned with OKEN Scanner

În sfîrșit, pentru orice variabilă aleatoare $\tilde{x} \in \mathfrak{L}^4(E, \mathcal{K}, P)$ se definește coeficientul de *exces* și se notează cu γ_2, coeficientul

$$\gamma_2 = \frac{V_4(\tilde{x})}{(V_2(\tilde{x}))^2} - 3. \tag{5.5.6}$$

5.6. CORELAȚIE. COEFICIENT DE CORELAȚIE

Fie $\tilde{x}$ și $\tilde{y}$ două variabile aleatoare pe cîmpul borelian de probabilitate $(E, \mathcal{K}, P)$. Dacă $\tilde{x}, \tilde{y} \in \mathfrak{L}^2(E, \mathcal{K}, P)$, numim *corelație (covariață)* a variabilelor aleatoare $\tilde{x}$ și $\tilde{y}$ și o notăm prin $\operatorname{cov}(\tilde{x}, \tilde{y})$, numărul real

$$\operatorname{cov}(\tilde{x}, \tilde{y}) = M[(\tilde{x} - M(\tilde{x}))(\tilde{y} - M(\tilde{y})]. \tag{5.6.1}$$

Din definiție rezultă imediat că avem $\operatorname{cov}(\tilde{x}, \tilde{x}) = V(\tilde{x})$. Se arată ușor că, dacă $\tilde{x}, \tilde{y}, \tilde{z} \in \mathfrak{L}^2(E, \mathcal{K}, P)$, atunci

$$\operatorname{cov}(\tilde{x}, \tilde{y}) = \operatorname{cov}(\tilde{y}, \tilde{x}), \tag{5.6.2}$$

$$\operatorname{cov}(\tilde{x}, \tilde{y}) = M(\tilde{x}\tilde{y}) - M(\tilde{x})\, M(\tilde{y}), \tag{5.6.3}$$

$$\operatorname{cov}(a\tilde{x}, b\tilde{y}) = ab\operatorname{cov}(\tilde{x}, \tilde{y}), \tag{5.6.4}$$

$$\operatorname{cov}(\tilde{x} + \tilde{y}, \tilde{z}) = \operatorname{cov}(\tilde{x}, \tilde{z}) + \operatorname{cov}(\tilde{y}, \tilde{z}), \tag{5.6.5}$$

unde a și b sînt numere reale arbitrare.

Variabilele aleatoare $\tilde{x}, \tilde{y} \in \mathfrak{L}^2(E, \mathcal{K}, P)$ sînt numite *necorelate* dacă $\operatorname{cov}(\tilde{x}, \tilde{y}) = 0$. Este clar că două variabile aleatoare independente sînt necorelate, după cum rezultă imediat din relația (5.6.3). Variabilele necorelate nu sînt totdeauna independente. Într-adevăr, variabilele aleatoare discrete $\tilde{x}$ și $\tilde{y}$, date prin repartițiile lor

$$\tilde{x}: \begin{pmatrix} -1 & 0 & 1 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}; \quad \tilde{y}: \begin{pmatrix} 0 & 1 \\ 1/3 & 2/3 \end{pmatrix},$$

sînt necorelate, deoarece $M(\tilde{x}) = 0$, $M(\tilde{y}) = 2/3$ și $M(\tilde{x}\tilde{y}) = 0$, ținînd seama că $\tilde{x}\tilde{y}$ are repartiția

$$\tilde{x}\tilde{y}: \begin{pmatrix} -1 & 0 & 1 \\ 2/9 & 5/9 & 2/9 \end{pmatrix}.$$

Scanned with OKEN Scanner

Aceste variabile nu sînt însă independente, deoarece avem, de exemplu,

$$P(\tilde{x} = -1,\ \tilde{y} = 0) = 0$$

și

$$P(\tilde{x} = -1)\ P(\tilde{y} = 0) = 1/9.$$

Numim *coeficient de corelație* al variabilelor aleatoare $\tilde{x},\ \tilde{y} \in \mathfrak{L}^2(E, \mathcal{K}, P)$, cu dispersii nenule, și îl notăm prin $r(\tilde{x}, \tilde{y})$, numărul real definit de relația

$$r(\tilde{x}, \tilde{y}) = \frac{\operatorname{cov}(\tilde{x}, \tilde{y})}{(V(\tilde{x})\ V(\tilde{y}))^{1/2}}. \tag{5.6.6}$$

Rezultă totodată că putem scrie

$$r(\tilde{x}, \tilde{y}) = \frac{M(\tilde{x}\tilde{y}) - M(\tilde{x})\,M(\tilde{y})}{(V(\tilde{x})\,V(\tilde{y}))^{1/2}}. \tag{5.6.7}$$

Proprietățile coeficientului de corelație sînt date în propoziția următoare.

Teorema 5.6.1. *Dacă* $\tilde{x},\ \tilde{y} \in \mathfrak{L}^2(E, \mathcal{K}, P)$ *și* $V(\tilde{x})V(\tilde{y}) \neq 0$, *atunci*

(a) Variabilele aleatoare $\tilde{x}$ *și* $\tilde{y}$ *sînt necorelate dacă și numai dacă* $r(\tilde{x},\tilde{y}) = 0$;

(b) $|r(\tilde{x}, \tilde{y})| \leqslant 1$;

(c) Dacă $|r(\tilde{x},\tilde{y})| = 1$, *atunci între* $\tilde{x}$ *și* $\tilde{y}$ *există o relație liniară.*

Demonstrație. (a) Rezultă imediat din faptul că $\operatorname{cov}(\tilde{x}, \tilde{y}) = 0$.

(b) Din inegalitatea lui Schwarz obținem

$$|M[(\tilde{x} - M(\tilde{x}))\ (\tilde{y} - M(\tilde{y}))]| \leqslant [M^2(\tilde{x} - M(\tilde{x}))\ M^2(\tilde{y} - M(\tilde{y}))]^{1/2},$$

de unde se obține imediat rezultatul dorit.

(c) Dacă notăm

$$\tilde{x}' = (\tilde{x} - M(\tilde{x}))/[V(\tilde{x})]^{1/2};\ \tilde{y}' = (\tilde{y} - M(\tilde{y}))/[V(\tilde{y})]^{1/2},$$

obținem $M_2(\tilde{x}') = M_2(\tilde{y}') = 1$ și $M(\tilde{x}'\tilde{y}') = r(\tilde{x}, \tilde{y}) = \pm 1$.

Scanned with OKEN Scanner

Rezultă atunci că

$$M((\tilde{x}' \pm \tilde{y}')^2) = M_2(\tilde{x}') + M_2(\tilde{y}') \pm 2\,M(\tilde{x}'\tilde{y}') = \\ = 2(1 \pm M(\tilde{x}'\tilde{y}'))$$

și deci avem fie $M((\tilde{x}' + \tilde{y}')^2) = 0$, fie $M((\tilde{x}' - \tilde{y}')^2) = 0$, adică $\tilde{x}' \pm \tilde{y}' = 0$ aproape sigur. Cu alte cuvinte, avem

$$\tilde{x} = M(\tilde{x}) - (\tilde{y} - M(\tilde{y}))\,(V(\tilde{x})/V(\tilde{y}))^{1/2},$$

ceea ce arată că între $\tilde{x}$ și $\tilde{y}$ există o relație liniară de forma $\tilde{x} = ay + b$ cu $a \neq 0$.

În plus, dacă $V(\tilde{x}) \neq 0$, atunci din definiție rezultă imediat

$$r(\tilde{x}, \tilde{x}) = 1;\ r(\tilde{x}, -\tilde{x}) = -1.$$

5.7. VALORI NUMERICE ASOCIATE UNUI VECTOR ALEATOR

Să considerăm un vector aleator* $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)'$ pe cîmpul borelian de probabilitate $(E, \mathcal{K}, P)$. Numim *media* vectorului $\tilde{\mathbf{x}}$ și o notăm prin $M(\tilde{\mathbf{x}})$ vectorul cu componentele $M(\tilde{x}_1), \ldots, M(\tilde{x}_n)$, adică $M(\tilde{\mathbf{x}}) = (M(\tilde{x}_1), \ldots, M(\tilde{x}_n))'$. Analog se poate defini media unei matrice cu elemente aleatoare. Mai exact, dacă $\tilde{\mathbf{X}} = (\tilde{x}_{ij})$, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$, este o matrice ale cărei elemente sînt variabile aleatoare, vom nota prin $M(\tilde{\mathbf{X}})$ matricea cu elementele $M(\tilde{x}_{ij})$, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$.

Numim *matrice de covarianță (corelație)* a vectorului aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)'$ și o notăm prin $V(\tilde{\mathbf{x}})$ sau prin $V(\tilde{x}_1, \ldots, \tilde{x}_n)$, matricea cu elementele $\operatorname{cov}(\tilde{x}_i, \tilde{x}_j)$, $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant n$.

Din definiție rezultă imediat că matricea de covarianță este o matrice simetrică, avînd pe diagonala principală dispersiile variabilelor aleatoare $\tilde{x}_1, \ldots, \tilde{x}_n$. Este de asemenea ușor de văzut că are loc formula

$$V(\tilde{\mathbf{x}}) = M[(\tilde{\mathbf{x}} - M(\tilde{\mathbf{x}}))\,(\tilde{\mathbf{x}} - M(\tilde{\mathbf{x}}))'] . \qquad (5.7.1)$$

---

* Accentul indică operația de transpunere a matricelor.

Dacă $\mathbf{A}$ este o matrice cu m linii și n coloane, iar $\tilde{\mathbf{y}} = \mathbf{A}\tilde{\mathbf{x}}$, atunci rezultă că

$$M(\tilde{\mathbf{y}}) = \mathbf{A}M(\tilde{\mathbf{x}}) \tag{5.7.2}$$

și

$$V(\tilde{\mathbf{y}}) = \mathbf{A}V(\tilde{\mathbf{x}})\mathbf{A}'. \tag{5.7.3}$$

Putem arăta acum că matricea de covarianță $V(\tilde{\mathbf{x}})$ este pozitiv semidefinită, adică

$$\mathbf{z}'V(\tilde{\mathbf{x}})\,\mathbf{z} \geqslant 0$$

pentru orice $\mathbf{z} \in R^n$. Într-adevăr, considerînd variabila aleatoare $\tilde{y} = \mathbf{z}'\tilde{\mathbf{x}}$ și ținînd seama de relațiile stabilite mai sus și de faptul că dispersia unei variabile aleatoare este nenegativă rezultă

$$V(\tilde{y}) = V(\mathbf{z}'\tilde{\mathbf{x}}) = \mathbf{z}'V(\tilde{\mathbf{x}})\mathbf{z}'.$$

Vom spune că vectorul aleator $\tilde{\mathbf{x}}$ este *nedegenerat (nesingular)* dacă $V(\tilde{\mathbf{x}})$ este o matrice nesingulară. Este clar deci că $\tilde{\mathbf{x}}$ este nesingular dacă și numai dacă $V(\tilde{\mathbf{x}})$ este pozitiv definită, adică dacă este pozitiv semidefinită și în plus avem $\mathbf{z}'V(\tilde{\mathbf{x}})\mathbf{z} = 0$ dacă și numai dacă $\mathbf{z} = \mathbf{0}$. Se poate arăta ușor că $\tilde{\mathbf{y}} = \mathbf{A}\tilde{\mathbf{x}}$ este nedegenerat dacă și numai dacă $\tilde{\mathbf{x}}$ este nedegenerat, $m \leqslant n$, iar $\operatorname{rang}(\mathbf{A}) = m$.

Uneori vom considera inversa matricei de covarianță $V(\tilde{\mathbf{x}})$, numită și *matricea de precizie* și notată prin $V^{-1}(\tilde{\mathbf{x}})$; evident, matricea de precizie se poate defini numai pentru vectori aleatori nedegenerați.

Dacă vectorul aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)'$ este partiționat în vectorul $\tilde{\mathbf{x}}_1 = (\tilde{x}_1, \ldots, \tilde{x}_m)'$ și $\tilde{\mathbf{x}}_2 = (\tilde{x}_{m+1}, \ldots, \tilde{x}_n)'$, atunci $M(\tilde{\mathbf{x}})$ și $V(\tilde{\mathbf{x}})$ pot fi partiționate corespunzător:

$$M(\tilde{\mathbf{x}}) = \begin{pmatrix} M(\tilde{\mathbf{x}}_1) \\ M(\tilde{\mathbf{x}}_2) \end{pmatrix}; \quad V(\tilde{\mathbf{x}}) = \begin{pmatrix} V(\tilde{\mathbf{x}}_1) & \mathbf{V}_{12} \\ \mathbf{V}_{21} & V(\tilde{\mathbf{x}}_2) \end{pmatrix},$$

unde $\mathbf{V}_{12}$ este matricea cu elementele $\operatorname{cov}(\tilde{x}_i, \tilde{x}_j)$, $1 \leqslant i \leqslant m$, $m+1 \leqslant j \leqslant n$, iar matricea $\mathbf{V}_{21}$ este matricea cu elementele $\operatorname{cov}(x_i, y_j)$, $m+1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant m$.

Scanned with OKEN Scanner

Noțiunea de moment se poate extinde și la vectori aleatori. Dacă $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)'$ este un vector aleator pe cîmpul borelian de probabilitate $(E, \mathfrak{K}, P)$, iar $(r_1, \ldots, r_n)' \in R^n$, atunci numim *moment de ordinul* $(r_1, \ldots, r_n)'$ al vectorului aleator $\tilde{\mathbf{x}}$ numărul real notat prin $M_{r_1 \ldots r_n}(\tilde{x}_1, \ldots, \tilde{x}_n)$ definit ca fiind media variabilei aleatoare $\tilde{x}_1^{r_1} \ldots \tilde{x}_n^{r_n}$, adică:

$$M_{r_1 \ldots r_n}(\tilde{x}_1, \ldots, \tilde{x}_n) = M(\tilde{x}_1^{r_1} \ldots \tilde{x}_n^{r_n}), \tag{5.7.4}$$

dacă această valoare medie există. Cînd nu există pericol de confuzie, vom scrie $M_{r_1 \ldots r_n}(\tilde{\mathbf{x}})$. Evident, avem

$$M_{r_1 \ldots r_n}(\tilde{\mathbf{x}}) = \int_E \tilde{x}_1^{r_1} \ldots \tilde{x}_n^{r_n} \, dP = \int_{R^n} x_1^{r_1} \ldots x_n^{r_n} \, dF(x_1, \ldots, x_n), \tag{5.7.5}$$

unde $F(x_1, \ldots, x_n)$ este funcția de repartiție a vectorului aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)'$.

Analog, *momentul centrat de ordinul* $(r_1, \ldots, r_n)'$ al vectorului aleator $\tilde{\mathbf{x}}$ este numărul real notat prin $V_{r_1 \ldots r_n}(\tilde{\mathbf{x}})$ sau $V_{r_1 \ldots r_n}(\tilde{x}_1, \ldots, \tilde{x}_n)$ și definit ca fiind media variabilei aleatoare $(\tilde{x}_1 - M(\tilde{x}_1))^{r_1} \ldots (\tilde{x}_n - M(\tilde{x}_n))^{r_n}$, adică

$$V_{r_1 \ldots r_n}(\tilde{\mathbf{x}}) = M[(\tilde{x}_1 - M(\tilde{x}_1))^{r_1} \ldots (\tilde{x}_n - M(\tilde{x}_n))^{r_n}], \tag{5.7.6}$$

dacă această valoare medie există. Evident,

$$V_{r_1 \ldots r_n}(\tilde{\mathbf{x}}) = \int_E (\tilde{x}_1 - M(\tilde{x}_1))^{r_1} \ldots (\tilde{x}_n - M(\tilde{x}_n))^{r_n} \, dP \tag{5.7.7}$$

și

$$V_{r_1 \ldots r_n}(\tilde{\mathbf{x}}) = \int_{R^n} (x_1 - M(\tilde{x}_1))^{r_1} \ldots (x_n - M(\tilde{x}_n))^{r_n} \, dF(x_1, \ldots, x_n). \tag{5.7.8}$$

Este clar acum că noțiunea de corelație (covarianță) a variabilelor $\tilde{x}$ și $\tilde{y}$ este un caz particular al noțiunii de moment centrat al unui vector aleator. Într-adevăr, momentul centrat de ordinul (1,1) al vectorului aleator $(\tilde{x}, \tilde{y})'$ este chiar covarianța variabilelor aleatoare $\tilde{x}$ și $\tilde{y}$, adică $V_{11}(\tilde{x}, \tilde{y}) = \operatorname{cov}(\tilde{x}, \tilde{y})$.

Scanned with OKEN Scanner

Capitolul 6

VALORI MEDII CONDIȚIONATE

6.1. CAZURI PARTICULARE

(a) Valoarea medie condiționată de un eveniment. Să considerăm mai întîi un cîmp borelian de probabilitate $(E, \mathcal{K}, P)$ și fie $A \in \mathcal{K}$ un eveniment de probabilitate nenulă. În § 3.5 am văzut că $(E, \mathcal{K}_A, P_A)$, în care

$$\mathcal{K}_A = \{A \cap B : B \in \mathcal{K}\},$$

iar P_A este probabilitatea condiționată de A, este un cîmp borelian de probabilitate.

Definiția 6.1.1. *Dacă* $\tilde{x}$ *este o variabilă aleatoare integrabilă pe* (E, $\mathcal{K}$, P), *atunci integrala Lebesgue a lui* $\tilde{x}$ *în raport cu probabilitatea* P_A *este numită media lui* $\tilde{x}$ *condiționată de* A *și este notată prin* $M(\tilde{x}|A)$.

Din această definiție rezultă că putem scrie

$$M(\tilde{x} \mid A) = \int_E \tilde{x}(\omega)\, dP_A(\omega) \tag{6.1.1}$$

sau

$$M(\tilde{x} \mid A) = \int_E \tilde{x}(\omega) P(d\omega \mid A). \tag{6.1.2}$$

Se poate arăta imediat că are loc și relația

$$M(\tilde{x} \mid A) = \frac{1}{P(A)} \int_A \tilde{x}(\omega)\, dP(\omega). \tag{6.1.3}$$

Scanned with OKEN Scanner

Într-adevăr, pentru o variabilă aleatoare simplă $\tilde{x}$ care ia valorile $x_1, \ldots, x_n$, avem

$$M(\tilde{x} \mid A) = \sum_{i=1}^{n} x_i P(\tilde{x} = x_i \mid A) = \frac{1}{P(A)} \sum_{i=1}^{n} x_i P(A \cap (\tilde{x} = x_i)),$$

de unde urmează imediat formula (6.1.3) pentru acest caz. Pentru o variabilă aleatoare oarecare $\tilde{x} \in \mathfrak{L}(E, \mathcal{K}, P)$ rezultatul se deduce din propoziția 4.1.4.

Dacă se consideră $B \in \mathcal{K}$, atunci

$$M(\chi_B \mid A) = \int_E \chi_B(\omega)\, dP_A(\omega) = P(B \mid A), \tag{6.1.4}$$

ceea ce arată că probabilitatea lui B condiționată de A este media indicatorului χ_B condiționată de A. În particular, pentru orice $x \in R$ și orice variabilă aleatoare $\tilde{x}$ pe $(E, \mathcal{K}, P)$ putem scrie

$$M(\chi_{\tilde{x}^{-1}((-\infty, x))} \mid A) = P(\tilde{x}^{-1}((-\infty, x)) \mid A) = P(\tilde{x} < x \mid A)). \tag{6.1.5}$$

Funcția de variabilă reală $F(x \mid A) = P(\tilde{x} < x \mid A)$, care se definește în acest mod, este numită *funcția de repartiție a lui $\tilde{x}$ condiționată de A*. Este clar că putem scrie

$$M(\tilde{x} \mid A) = \int_R x\, dF(x \mid A). \tag{6.1.6}$$

Dacă funcția de repartiție condiționată $F(x|A)$ are o densitate de probabilitate, aceasta va fi numită *densitatea de probabilitate a lui $\tilde{x}$ condiționată de A* și va fi notată prin $f(x|A)$. Evident, în acest caz, vom avea

$$M(\tilde{x} \mid A) = \int_R x f(x|A)\, dx. \tag{6.1.7}$$

În sfîrșit, dacă $\tilde{x}$ este discretă avînd valorile $(x_i)_{i \in I}$ putem considera funcția de masă condiționată $f(x \mid A)$ definită ca fiind egală cu $P(\tilde{x} = x_i \mid A)$ pentru $x = x_i$ și nulă în caz contrar.

Valoarea medie condiționată de A se poate atunci scrie în forma

$$M(\tilde{x} \mid A) = \sum_{i \in I} x_i f(x_i \mid A). \tag{6.1.8}$$

(b) Valoarea medie condiționată de o desfacere. Fie $(E, \mathcal{K}, P)$ un cîmp borelian de probabilitate, o variabilă aleatoare $\tilde{x} \in \mathfrak{L}(E, \mathcal{K}, P)$ și o desfacere $\mathcal{D} = (A_i)_{i \in I} \subset \mathcal{K}$.

Definiția 6.1.2. *Numim valoare medie a lui* $\tilde{x}$ *condiționată de desfacerea* $\mathcal{D}$ *și o notăm prin* $M(\tilde{x} \mid \mathcal{D})$, *funcția definită pe* E *și cu valori în* R *dată prin relația*

$$M(\tilde{x} \mid \mathcal{D}) = M(\tilde{x} \mid A_i)$$

pentru orice $\omega \in A_i$, *dacă* $P(A_i) > 0$, *și luînd valori arbitrare în rest.*

Este clar că funcția $M(\tilde{x} \mid \mathcal{D})$ astfel definită este o variabilă aleatoare pe $(E, \mathcal{K}, P)$ și este determinată abstracție făcînd de evenimente de probabilitate nulă.

Dacă $\mathfrak{B}(\mathcal{D})$ este corpul borelian generat de desfacerea $\mathcal{D}$, atunci putem calcula valoarea medie a unei variabile aleatoare $\tilde{x} \in \mathfrak{L}(E, \mathcal{K}, P)$ condiționată de un eveniment $A \in \mathfrak{B}(\mathcal{D})$ de probabilitate nenulă. Într-adevăr, dacă A se scrie sub forma

$$A = \bigcup_{j \in N} A_{i_j},$$

atunci rezultă

$$M(\tilde{x} \mid A) = \frac{1}{P(A)} \sum_{j \in N} M(\tilde{x} \mid A_{i_j}) P(A_{i_j}),$$

relație care se mai scrie și sub forma

$$\int_A \tilde{x}(\omega)\, dP(\omega) = \int_A M(\tilde{x} \mid \mathcal{D})\, dP. \tag{6.1.9}$$

Vom vedea în cele ce urmează că relația (6.1.9) este o relație ce caracterizează valoarea medie condiționată.

Scanned with OKEN Scanner

Dacă $B \in \mathcal{K}$, atunci media indicatorului său χ_B condiționată de desfacerea $\mathcal{D}$ este numită *probabilitatea lui* B *condiționată de desfacerea* $\mathcal{D}$ și este notată prin $P(B \mid \mathcal{D})$. Din relația (6.1.9) rezultă că

$$P(A \cap B) = \int_A P(B \mid \mathcal{D})\, dP, \tag{6.1.10}$$

relație care se obține pentru $\tilde{x} = \chi_B$, unde $A \in \mathcal{B}(\mathcal{D})$. Pentru $A = E$, obținem deci

$$P(B) = \int_E P(B \mid \mathcal{D})\, dP. \tag{6.1.11}$$

În particular, dacă $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)'$ este un vector aleator pe $(E, \mathcal{K}, P)$, atunci $P(\tilde{x}_1 < x_1, \ldots, \tilde{x}_n < x_n \mid \mathcal{D})$ este numită *funcția de repartiție a vectorului aleator* $\tilde{x}$ *condiționată de desfacerea* $\mathcal{D}$ și este notată prin $F(x_1, \ldots, x_n \mid \mathcal{D})$. Evident, din relația (6.1.11) se deduce ușor că

$$F(x_1, \ldots, x_n) = \int_E F(x_1, \ldots, x_n \mid \mathcal{D})\, dP(\omega), \tag{6.1.12}$$

unde F este funcția de repartiție a vectorului aleator $\tilde{\mathbf{x}}$.

(c) Valoarea medie condiționată de o variabilă aleatoare discretă. Să considerăm cîmpul borelian de probabilitate $(E, \mathcal{K}, P)$, o variabilă aleatoare $\tilde{x} \in \mathcal{L}(E, \mathcal{K}, P)$ și o variabilă aleatoare discretă $\tilde{y}$ pe $(E, \mathcal{K}, P)$ cu mulțimea valorilor $(y_i)_{i \in I}$. Dacă notăm

$$A_i = \{\omega : \tilde{y}(\omega) = y_i\}, \quad i \in I,$$

atunci este clar că $\mathcal{D} = (A_i)_{i \in I}$ este o desfacere.

Definiția 6.1.3. *Numim valoare medie a lui* $\tilde{x}$ *condiționată de variabila aleatoare discretă* $\tilde{y}$ *și o notăm prin* $M(\tilde{x} \mid \tilde{y})$, *valoarea medie a lui* $\tilde{x}$ *condiționată de desfacerea* $\mathcal{D} = (A_i)_{i \in I}$ *definită de* $\tilde{y}$.

Din definiție rezultă deci că dacă $P(\tilde{y} = y_i) > 0$, atunci

$$M(\tilde{x} \mid \tilde{y})(\omega) = M(\tilde{x} \mid \tilde{y}^{-1}(\{y_i\}))$$

pentru orice $\omega \in \tilde{y}^{-1}(\{y_i\})$, funcția $M(\tilde{x} \mid \tilde{y})$ fiind definită arbitrar în rest.

Dacă se consideră un eveniment

$$A = \bigcup_{i \in J} \{\omega : \tilde{y}(\omega) = y_i\}$$

unde $J \subset I$, atunci $A \in \mathfrak{B}(\mathfrak{D})$ și deci din (6.1.9) rezultă

$$\int_A \tilde{x}(\omega)\, dP(\omega) = \int_A M(\tilde{x} \mid \tilde{y})\, dP. \tag{6.1.13}$$

Ca și în cazul precedent, putem defini pentru $B \in \mathfrak{K}$ *probabilitatea lui B condiționată de* $\tilde{y}$, pe care o vom nota prin $P(B \mid \tilde{y})$. Prin urmare avem prin definiție

$$P(B \mid \tilde{y}) = M(\chi_B \mid \tilde{y}) \tag{6.1.14}$$

și deci

$$P(A \cap B) = \int_A P(B \mid \tilde{y})\, dP \tag{6.1.15}$$

pentru A definit mai sus. Evident, pentru $A = E$, se obține

$$P(B) = \int_E P(B \mid \tilde{y})\, dP. \tag{6.1.16}$$

În particular, dacă $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)'$ este un vector aleator pe $(E, \mathfrak{K}, P)$, atunci $P(\tilde{x}_1 < x_1, \ldots, \tilde{x}_n < x_n \mid \tilde{y})$ este numită *funcția de repartiție a vectorului aleator* $\tilde{\mathbf{x}}$ *condiționată de* $\tilde{y}$ și este notată prin $F(x_1, \ldots, x_n \mid \tilde{y})$. Evident, din (6.1.16) urmează

$$F(x_1, \ldots, x_n) = \int_E F(x_1, \ldots, x_n \mid \tilde{y})\, dP(\omega), \tag{6.1.17}$$

unde $F(x_1, \ldots, x_n)$ este funcția de repartiție a vectorului aleator $\tilde{\mathbf{x}}$.

Scanned with OKEN Scanner

6.2. VALOAREA MEDIE CONDIȚIONATĂ DE UN CORP BORELIAN

Să considerăm un cîmp borelian de probabilitate $(E, \mathcal{K}, P)$ și fie $\mathcal{F} \subset \mathcal{K}$ un corp borelian pe E. Dacă P' este restricția lui P la $\mathcal{F}$, atunci este clar că $(E, \mathcal{F}, P')$ este de asemenea un cîmp borelian de probabilitate.

Dacă $\tilde{x} \in \mathcal{L}(E, \mathcal{K}, P)$, atunci se poate arăta că există o variabilă aleatoare $\tilde{y}$ pe $(E, \mathcal{F}, P')$ cu proprietățile:

(a) $\tilde{y}$ *este integrabilă pe* $(E, \mathcal{F}, P')$;

$$(b) \quad \int_A \tilde{y}(\omega)\, dP(\omega) = \int_A \tilde{x}(\omega)\, dP(\omega)$$

pentru orice $A \in \mathcal{F}$. Existența variabilei aleatoare $\tilde{y}$ cu proprietățile *(a)* și *(b)* de mai sus rezultă simplu din binecunoscuta teoremă a lui Radon-Nicodym, pe care n-o demonstrăm însă aici (vezi, de exemplu, Cuculescu, I. (1976)).

Se poate de asemenea arăta ușor că, dacă $\tilde{z}$ este o altă variabilă aleatoare pe $(E, \mathcal{F}, P)$ care are proprietățile *(a)* și *(b)* de mai sus, atunci $\tilde{y} = \tilde{z}$ aproape sigur în raport cu P'.

Definiția 6.2.1. *Numim valoare medie a variabilei aleatoare* $\tilde{x}$ *condiționată de corpul borelian* $\mathcal{F}$ *și o notăm prin* $M(\tilde{x} \mid \mathcal{F})$, *una dintre variabilele aleatoare* $\tilde{y}$ *pe* $(E, \mathcal{F}, P')$ *care are proprietățile* (a) *și* (b) *de mai sus.*

Cu alte cuvinte, $M(\tilde{x} \mid \mathcal{F})$ este o variabilă aleatoare integrabilă pe $(E, \mathcal{F}, P')$ caracterizată de proprietatea

$$\int_A M(\tilde{x} \mid \mathcal{F})\, dP = \int_A \tilde{x}\, dP \qquad (6.2.1)$$

pentru orice $A \in \mathcal{F}$.

În particular, *probabilitatea unui eveniment* $B \in \mathcal{K}$ *condiționată de corpul borelian* $\mathcal{F}$ este notată prin $P(B \mid \mathcal{F})$ și este definită de relația

$$P(B \mid \mathcal{F}) = M(\chi_B \mid \mathcal{F}). \qquad (6.2.2)$$

Scanned with OKEN Scanner

Exemple. 1. Să considerăm corpul borelian banal $\mathcal{K}_0 = \{\varnothing, E\}$. În acest caz $M(\tilde{x} \mid \mathcal{K}_0) = M(\tilde{x})$, cum rezultă ușor din caracterizarea (6.2.1) a valorii medii condiționate.

2. Fie $A \in \mathcal{K}$ pentru care $0 < P(A) < 1$ și fie

$$\mathcal{K}_A = \{\varnothing, A, A^c, E\}.$$

Rezultă imediat că

$$M(\tilde{x} \mid \mathcal{K}_A) = \begin{cases} M(\tilde{x} \mid A) & \text{pentru } \omega \in A, \\ M(\tilde{x} \mid A^c) & \text{pentru } \omega \in A^c. \end{cases}$$

3. Fie $\mathcal{D} = (A_i)_{i \in I} \subset \mathcal{K}$ o desfacere și fie $\mathcal{B}(\mathcal{D})$ corpul borelian generat de $\mathcal{D}$. Ținînd seama de caracterizarea (6.2.1) și de proprietatea (6.1.9) a valorii medii condiționate de desfacerea $\mathcal{D}$ rezultă imediat că

$$M(\tilde{x} \mid \mathcal{B}(\mathcal{D})) = M(\tilde{x} \mid \mathcal{D}).$$

4. Fie $\tilde{y}$ o variabilă aleatoare discretă pe $(E, \mathcal{K}, P)$ și fie $\mathcal{B}(\tilde{y})$ corpul borelian generat de această variabilă aleatoare. Din caracterizarea (6.2.1) pentru $M(\tilde{x} \mid \mathcal{B}(\tilde{y}))$ și din proprietatea (6.1.13) a lui $M(\tilde{x} \mid \tilde{y})$ rezultă că

$$M(\tilde{x} \mid \mathcal{B}(\tilde{y})) = M(\tilde{x} \mid \tilde{y}).$$

Proprietățile valorii medii condiționate sînt date în teorema următoare. O parte dintre afirmațiile și relațiile prezentate în teoremă au loc aproape sigur. Deoarece din context rezultă clar acest lucru, nu considerăm necesar să precizăm aceasta de fiecare dată.

Teorema 6.2.1. *Fie* $(E, \mathcal{K}, P)$ *un cîmp borelian de probabilitate,* $\mathcal{F} \subset \mathcal{K}$ *un corp borelian pe* E, *iar* P' *restricția la* $\mathcal{F}$ *a lui* P. *În aceste condiții avem:*

(a) Dacă $\tilde{x}$ *este* $\mathcal{F}$*-măsurabilă, atunci*

$$M(\tilde{x} \mid \mathcal{F}) = \tilde{x}; \tag{6.2.3}$$

(b) Dacă $\tilde{x} \geqslant 0$ *și* $\tilde{x} \in \mathcal{L}(E, \mathcal{K}, P)$, *atunci* $M(\tilde{x} \mid \mathcal{F}) \geqslant 0$;

(c) Dacă $\tilde{x} \leqslant \tilde{y}$ *și* $\tilde{x}, \tilde{y} \in \mathcal{L}(E, \mathcal{K}, P)$, *atunci*

$$M(\tilde{x} \mid \mathcal{F}) \leqslant M(\tilde{y} \mid \mathcal{F}); \tag{6.2.4}$$

(d) Dacă $\tilde{x} \in \mathfrak{L}(E, \mathcal{K}, P)$, *atunci*

$$M(M(\tilde{x} \mid \mathcal{F})) = M(\tilde{x}); \tag{6.2.5}$$

(e) Dacă a_1 *și* a_2 *sînt numere reale, iar* $\tilde{x}_1$ *și* $\tilde{x}_2$ *sînt integrabile pe* $(E, \mathcal{K}, P)$, *atunci*

$$M(a_1\tilde{x}_1 + a_2\tilde{x}_2 \mid \mathcal{F}) = a_1M(\tilde{x}_1 \mid \mathcal{F}) + a_2M(\tilde{x}_2 \mid \mathcal{F}); \tag{6.2.6}$$

(f) Dacă $(\tilde{x}_n)$ *este un șir crescător de variabile aleatoare nenegative și* $\lim_n \tilde{x}_n$ *este integrabilă, atunci*

$$\lim_n M(\tilde{x}_n \mid \mathcal{F}) = M(\lim_n \tilde{x}_n \mid \mathcal{F}), \tag{6.2.7}$$

șirul $M(\tilde{x}_n \mid \mathcal{F})$ *fiind de asemenea crescător;*

(g) Dacă $\tilde{y}$ *este* $\mathcal{F}$*-măsurabilă, iar* $\tilde{x}$ *și* $\tilde{x}\tilde{y}$ *sînt integrabile pe* $(E, \mathcal{F}, P')$, *atunci*

$$M(\tilde{x}\tilde{y} \mid \mathcal{F}) = \tilde{y}M(\tilde{x} \mid \mathcal{F}); \tag{6.2.8}$$

(h) Dacă $\mathcal{F}' \subset \mathcal{F}$, *atunci*

$$M(M(\tilde{x} \mid \mathcal{F}) \mid \mathcal{F}') = M(\tilde{x} \mid \mathcal{F}'). \tag{6.2.9}$$

Demonstrație. (a) și *(b)* rezultă imediat din definiția valorii medii condiționate.

(c) Deoarece $\tilde{y} - \tilde{x} \geqslant 0$, afirmația rezultă din *(a)* și *(b)*.

(d) Se obține imediat din definiție.

(e) Să arătăm că $a_1M(\tilde{x}_1 \mid \mathcal{F}) + a_2M(\tilde{x}_2 \mid \mathcal{F})$ satisface caracterizarea lui $M(a_1\tilde{x}_1 + a_2\tilde{x}_2 \mid \mathcal{F})$:

$$\int_A [a_1M(\tilde{x}_1|\mathcal{F}) + a_2M(\tilde{x}_2|\mathcal{F})]\,dP = a_1\int_A M(\tilde{x}_1|\mathcal{F})\,dP + a_2\int_A M(\tilde{x}_2|\mathcal{F})\,dP =$$

$$= a_1\int_A \tilde{x}_1\,dP + a_2\int_A \tilde{x}_2\,dP = \int_A (a_1\tilde{x}_1 + a_2\tilde{x}_2)\,dP,$$

pentru orice $A \in \mathcal{F}$.

(f) Din *(c)* rezultă că șirul $M(\tilde{x}_n \mid \mathcal{F})$ este un șir crescător de funcții $\mathcal{F}$-măsurabile nenegative. Aplicînd proprietatea de convergență monotonă a integralei Lebesgue, rezultă

$$\int_A [\lim_n M(\tilde{x}_n \mid \mathcal{F})]\, dP = \lim_n \int_A M(\tilde{x}_n \mid \mathcal{F})\, dP =$$

$$= \lim_n \int_A \tilde{x}_n\, dP = \int_A (\lim_n \tilde{x}_n)\, dP$$

pentru orice $A \in \mathcal{F}$.

Pentru $A = E$ rezultă deci că $\lim_n M(\tilde{x}_n \mid \mathcal{F})$ este o variabilă aleatoare integrabilă și rezultatul căutat se deduce din relația stabilită mai sus.

(g) Evident, ne putem limita la cazul $\tilde{x} \geqslant 0$, $\tilde{y} \geqslant 0$. Dacă $A \in \mathcal{F}$ și $\tilde{y} = \chi_A$, atunci

$$\int_A \tilde{y} M(\tilde{x} \mid \mathcal{F})\, dP = \int_{A \cap B} M(\tilde{x} \mid \mathcal{F})\, dP = \int_{A \cap B} \tilde{x}\, dP = \int_A \tilde{x}\tilde{y}\, dP,$$

ceea ce arată că relația (6.2.8) este adevărată pentru $\tilde{y}$ aparținînd clasei indicatorilor evenimentelor din $\mathcal{F}$. Folosind faptul că $\tilde{y} \geqslant 0$ este limita unui șir crescător de variabile aleatoare simple și utilizînd proprietățile *(e)* și *(f)*, relația (6.2.8) se obține pentru orice $\tilde{x} \geqslant 0$ și orice $\tilde{y} \geqslant 0$.

(h) Deoarece $M(\tilde{x} \mid \mathcal{F})$ este o funcție $\mathcal{F}$-măsurabilă, rezultă că pentru $\mathcal{F}' \subset \mathcal{F}$ putem defini media variabilei aleatoare $M(\tilde{x} \mid \mathcal{F})$ condiționată de $\mathcal{F}'$, adică $M(M(\tilde{x} \mid \mathcal{F}) \mid \mathcal{F}')$. Pentru a demonstra rezultatul (6.2.9), fie $A \in \mathcal{F}'$ și deci $A \in \mathcal{F}$. Avem

$$\int_A M(M(\tilde{x} \mid \mathcal{F}) \mid \mathcal{F}')\, dP = \int_A M(\tilde{x} \mid \mathcal{F})\, dP = \int_A \tilde{x}\, dP = \int_A M(\tilde{x} \mid \mathcal{F})\, dP,$$

de unde rezultatul căutat.

Scanned with OKEN Scanner

Observații. Din demonstrația ultimei proprietăți rezultă de asemenea că dacă $\tilde{y}$ este $\mathcal{F}$-măsurabilă, atunci

$$M(\tilde{x}\tilde{y} \mid \mathcal{F}') = M(\tilde{y}M(\tilde{x} \mid \mathcal{F}) \mid \mathcal{F}'), \tag{6.2.10}$$

cu condiția ca $\tilde{x}$ și $\tilde{x}\tilde{y}$ să fie integrabile.

Evident, unele dintre proprietățile valorii medii condiționate se pot transcrie într-o formă interesantă pentru probabilitatea condiționată de un corp borelian. Se obțin astfel următoarele proprietăți:

(a′) Dacă $A \in \mathcal{F}$, *atunci*

$$P(A \mid \mathcal{F}) = \chi_A; \tag{6.2.11}$$

(c′) Dacă $A \subset B$, *cu* $A, B \in \mathcal{F}$, *atunci*

$$P(A \mid \mathcal{F}) \leqslant P(B \mid \mathcal{F}); \tag{6.2.12}$$

(d′) Dacă $A \in \mathcal{F}$, *atunci*

$$M(P(A \mid \mathcal{F})) = P(A); \tag{6.2.13}$$

(e′) Dacă $A, B \in \mathcal{F}$ *și* $A \cap B = \varnothing$, *atunci*

$$P(A \cup B \mid \mathcal{F}) = P(A \mid \mathcal{F}) + P(B \mid \mathcal{F}); \tag{6.2.14}$$

(f′) Dacă $(A_n) \subset \mathcal{F}$ *este un șir crescător de evenimente, atunci*

$$\lim_n P(A_n \mid \mathcal{F}) = P(\bigcup_n A_n \mid \mathcal{F}). \tag{6.2.15}$$

Dacă, în plus, $A_i \cap A_j = \varnothing$ *pentru* $i \neq j$, *atunci*

$$P(\bigcup_{n \in N} A_n \mid \mathcal{F}) = \sum_{n \in N} P(A_n \mid \mathcal{F}). \tag{6.2.16}$$

Să observăm că proprietatea (6.2.14) stabilește aditivitatea simplă a probabilității condiționate de un corp borelian. La prima vedere s-ar părea că proprietatea (6.2.16) stabilește și aditivitatea completă a acestei probabilități condiționate. În general, acest lucru poate să nu fie adevărat. Într-adevăr, pentru un șir $(A_n) \subset \mathcal{F}$ dat cu proprietatea că $A_i \cap A_j = \varnothing$ pentru $i \neq j$, relația (6.2.16) are loc aproape sigur, dar evenimentele de probabilitate nulă pentru care (6.2.16) nu are loc depind de alegerea șirului (A_n). Este deci posibil să nu existe $\omega \in E$ pentru care (6.2.16) să fie adevărată pentru toate șirurile $(A_n) \subset \mathcal{F}$ cu proprietatea că $A_i \cap A_j = \varnothing$ pentru $i \neq j$.

În sfîrșit, să observăm că, pentru orice $A \in \mathcal{F}$, avem

$$0 \leqslant P(A \mid \mathcal{F}) \leqslant 1 \tag{6.2.17}$$

și

$$P(E \mid \mathcal{F}) = 1. \tag{6.2.18}$$

Într-adevăr, dacă ținem seama de *(a)* și *(b)*, prima relație decurge imediat din faptul că $0 \leqslant \chi_A \leqslant 1$. A doua relație este o consecință imediată a proprietății (6.2.11).

6.3. VALOAREA MEDIE CONDIȚIONATĂ DE O VARIABILĂ ALEATOARE

Am văzut în § 6.2 cum se poate defini valoarea medie condiționată de o variabilă aleatoare discretă. În cele ce urmează vom defini valoarea medie condiționată de o variabilă aleatoare oarecare $\tilde{y}$ pe $(E, \mathcal{K}, P)$.

Definiția 6.3.1. *Numim valoare medie a variabilei aleatoare* $\tilde{x} \in \mathcal{L}(E, \mathcal{K}, P)$ *condiționată de variabila aleatoare* $\tilde{y}$ *și o notăm prin* $M(\tilde{x} \mid \tilde{y})$, *valoarea medie a lui* $\tilde{x}$ *condiționată de corpul borelian** $\mathcal{B}(\tilde{y})$ *generat de* $\tilde{y}$.

Avem deci prin definiție

$$M(\tilde{x} \mid \tilde{y}) = M(\tilde{x} \mid \mathcal{B}(\tilde{y})). \tag{6.3.1}$$

Rezultă de asemenea că valoarea medie condiționată de o variabilă aleatoare este caracterizată de proprietatea

$$\int_{\tilde{y}^{-1}(B)} M(\tilde{x} \mid \tilde{y})\, dP = \int_{\tilde{y}^{-1}(B)} \tilde{x}\, dP, \tag{6.3.2}$$

valabilă pentru orice $B \in \mathcal{B}$.

---

* Vezi § 4.1 pentru definiția lui $\mathcal{B}(\tilde{y})$.

În particular, dacă $A \in \mathcal{K}$, definim probabilitatea lui A condiționată de variabila aleatoare $\tilde{y}$ și o notăm prin $P(A \mid \tilde{y})$ cu ajutorul relației

$$P(A \mid \tilde{y}) = M(\chi_A \mid \tilde{y}). \tag{6.3.3}$$

Din teorema lui Radon-Nicodym urmează că $M(\tilde{x} \mid \tilde{y})$ este o funcție măsurabilă de $\tilde{y}$. Cu alte cuvinte, există o funcție $g : R \to R$, măsurabilă, cu proprietatea că

$$M(\tilde{x} \mid \tilde{y}) = g(\tilde{y}). \tag{6.3.4}$$

Funcția g cu această proprietate este definită în mod unic de relația

$$\int_{\tilde{y}^{-1}(B)} \tilde{x}\, dP = \int_B g(x)\, dP_{\tilde{x}}(x), \tag{6.3.5}$$

unde $P_{\tilde{x}}(A) = P(\tilde{x}^{-1}(A))$ pentru orice $A \in \mathcal{B}$.

Din proprietățile valorii medii condiționate de un corp borelian, care au fost demonstrate în § 6.2, rezultă următoarele proprietăți ale valorii medii condiționate de o variabilă aleatoare:

(a) Dacă $\tilde{x}$ *este* $\mathcal{B}(\tilde{y})$*-măsurabilă, atunci*

$$M(\tilde{x} \mid \tilde{y}) = \tilde{x}. \tag{6.3.6}$$

În particular, avem

$$M(\tilde{y} \mid \tilde{y}) = \tilde{y}; \tag{6.3.7}$$

(b) Dacă $\tilde{x}$ *este integrabilă pe* $(E, \mathcal{K}, P)$, *atunci*

$$M(M(\tilde{x} \mid \tilde{y})) = M(\tilde{x}); \tag{6.3.8}$$

(c) Dacă $\tilde{x}_1$ *este* $B(\tilde{y})$*-măsurabilă, iar* $\tilde{x}_2$ *și* $\tilde{x}_1\tilde{x}_2$ *sînt integrabile pe* $(E, \mathcal{K}, P)$, *atunci*

$$M(\tilde{x}_1\tilde{x}_2 \mid \tilde{y}) = \tilde{x}_1\, M(\tilde{x}_2 \mid \tilde{y}). \tag{6.3.9}$$

În particular, avem

$$M(\tilde{x}_1\tilde{x}_2 \mid \tilde{x}_1) = \tilde{x}_1 M(\tilde{x}_2 \mid \tilde{x}_1). \tag{6.3.10}$$

Scanned with OKEN Scanner

În sfîrșit, să observăm că, dacă $\tilde{x}$ și $\tilde{y}$ sînt independente iar $\tilde{x}$ este integrabilă pe $(E, \mathcal{K}, P)$, atunci

$$M(\tilde{x} \mid \tilde{y}) = M(\tilde{x}). \tag{6.3.11}$$

Într-adevăr, din independența variabilelor aleatoare $\tilde{x}$ și $\tilde{y}$ rezultă

$$\int_{\tilde{y}^{-1}(B)} \tilde{x}\, dP = M(\tilde{x}\chi_{\tilde{y}^{-1}(B)}) = M(\tilde{x})\, P_{\tilde{y}}(B),$$

pentru orice $B \in \mathcal{B}$, ceea ce arată că $M(\tilde{x})$ satisface caracterizarea (6.3.2) a lui $M(\tilde{x} \mid \tilde{y})$.

În particular, dacă $A \in \mathcal{K}$ și $\tilde{y}$ sînt independente, atunci

$$P(A \mid \tilde{y}) = P(A). \tag{6.3.12}$$

Să observăm în încheiere că se poate defini valoarea medie condiționată de o familie de variabile aleatoare $(\tilde{y}_i)_{i \in I}$ pe $(E, \mathcal{K}, P)$.

Definiția 6.3.2. *Dacă* $\tilde{x}$ *este integrabilă pe* (E, $\mathcal{K}$, P), *numim valoarea medie a lui* $\tilde{x}$ *condiționată de variabilele aleatoare* $(\tilde{y}_i)_{i \in I}$ *și o notăm prin* $M(\tilde{x} \mid (\tilde{y}_i)_{i \in I})$, *valoarea medie a lui* $\tilde{x}$ *condiționată de corpul borelian generat de această familie* $\mathcal{B}(\tilde{y}_i \mid i \in I)$.

Avem deci prin definiție

$$M(\tilde{x} \mid (\tilde{y}_i)_{i \in I}) = M(\tilde{x} \mid \mathcal{B}(\tilde{y}_i \mid i \in I)). \tag{6.3.13}$$

În cazul unei familii finite de variabile aleatoare $\tilde{y}_1, \ldots, \tilde{y}_n$, valoarea medie a lui $\tilde{x}$ condiționată de această familie va fi notată mai simplu prin $M(\tilde{x} \mid \tilde{y}_1, \ldots, \tilde{y}_n)$.

În legătură cu această definiție este de semnalat proprietatea următoare, care se obține imediat din (6.2.9): dacă $(\tilde{y}_1, \tilde{y}_2)'$ este un vector aleator pe $(E, \mathcal{K}, P)$, iar $\tilde{x}$ este integrabilă, atunci

$$M(M(\tilde{x} \mid (\tilde{y}_1, \tilde{y}_2)) \mid \tilde{y}_1) = M(\tilde{x} \mid \tilde{y}_1). \tag{6.3.14}$$

Scanned with OKEN Scanner

6.4. FUNCȚII DE REPARTIȚIE CONDIȚIONATE

În § 6.1 am prezentat noțiunea de funcție de repartiție condiționată de un eveniment de probabilitate nenulă, cu ajutorul căreia se poate calcula valoarea medie condiționată de acest eveniment (vezi formula (6.1.5)). Problema care se pune este de a vedea dacă se poate introduce o noțiune analoagă pentru cazul general.

Fie $(E, \mathcal{K}, P)$ un cîmp borelian de probabilitate și $\mathcal{F} \subset \mathcal{K}$ un corp borelian pe E. Așa cum am observat mai sus, $P(A \mid \mathcal{F})$, care este definită pentru fiecare $\omega \in E$ și $A \in \mathcal{K}$, nu poate fi considerată în general ca o probabilitate pe $\mathcal{K}$ pentru un $\omega \in E$ fixat. În legătură cu această situație introducem noțiunea de *probabilitate regulată*.

Definiția 6.4.1. *Probabilitatea condiționată de un corp borelian* $\mathrm{P}(\mathrm{A} \mid \mathcal{F})$ *este numită regulată dacă există o funcție* $\mathrm{q} : \mathrm{E} \times \mathcal{K} \to \mathrm{R}$ *cu proprietățile următoare:*

(*a*) *Pentru orice* $\omega \in \mathrm{E}$ *fixat,* $\mathrm{q}(\omega, \cdot)$ *este o probabilitate pe* $\mathcal{K}$*;*

(*b*) *Pentru orice* $\mathrm{A} \in \mathcal{K}$ *fixat avem*

$$\mathrm{P}(\mathrm{A} \mid \mathcal{F})\,(\omega) = \mathrm{q}(\omega, \mathrm{A}),$$

aproape sigur în raport cu P.

Dacă $P(A \mid \mathcal{F})$ este regulată, atunci familia $(P(A \mid \mathcal{F}))_{\omega \in E}$ se identifică cu $q(\omega, A)$.

Teorema 6.4.1. *Dacă* $\tilde{\mathrm{x}}$ *este o variabilă aleatoare integrabilă pe* $(\mathrm{E}, \mathcal{K}, \mathrm{P})$, $\mathcal{F} \subset \mathcal{K}$ *este un corp borelian pe* E, *iar probabilitatea condiționată de* $\mathcal{F}$ *este regulată, atunci*

$$\mathrm{M}(\tilde{\mathrm{x}} \mid \mathcal{F}) = \int_E \tilde{\mathrm{x}}(\omega)\, \mathrm{P}(\mathrm{d}\omega \mid \mathcal{F}) \tag{6.4.1}$$

aproape sigur în raport cu P.

Demonstrație. Dacă $A \in \mathcal{K}$, atunci (6.4.1) este adevărată pentru $\tilde{x}(\omega) = \chi_A(\omega)$ conform definiției celor doi termeni. Relația rămîne adevărată pentru o variabilă aleatoare simplă $\tilde{x}$, ținînd seama de faptul că cei doi termeni din (6.4.1) au proprietatea de liniaritate ca funcționale de $\tilde{x}$. Considerînd șiruri crescătoare de variabile aleatoare simple și trecînd la limită în (6.4.1), se obține rezultatul pentru variabile aleatoare nenega-

Scanned with OKEN Scanner

tive. Cazul general se obține prin procedeul obișnuit, utilizînd din nou liniaritatea celor doi termeni din (6.4.1).

Rezultatul precedent ne arată că proprietatea (6.1.5) este adevărată în cazul general, cu condiția ca probabilitatea condiționată de corpul borelian să fie regulată.

Deoarece proprietatea de regularitate a probabilității condiționate nu este totdeauna adevărată, sîntem conduși la lărgirea acestei noțiuni.

Fie $(E, \mathcal{K}, P)$ un cîmp borelian de probabilitate, $\mathcal{F} \subset \mathcal{K}$ un corp borelian pe E și $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)'$ un vector aleator pe $(E, \mathcal{K}, P)$.

Definiția 6.4.2. *Funcția* $Q : E \times \mathcal{B}^n \to R$ *este numită repartiție regulată condiționată a vectorului aleator* $\tilde{\mathbf{x}}$ *relativ la corpul borelian* $\mathcal{F}$ *dacă îndeplinește condițiile următoare:*

(*a*) *Pentru orice* $B \in \mathcal{B}^n$ *avem*

$$Q(\omega, B) = P(\tilde{\mathbf{x}}^{-1}(B) \mid \mathcal{F}) \qquad (6..4.2)$$

aproape sigur în raport cu P.

(*b*) *Pentru orice* $B \in \mathcal{B}^n$, *funcția* $Q(\cdot, B)$ *este* $\mathcal{F}$*-măsurabilă;*

(*c*) $P(\{\omega : Q(\omega, \cdot)$ *este probabilitate pe* $\mathcal{B}^n\}) = 1.$ (6.4.3)

Se poate dovedi ușor că (6.4.2) este echivalentă cu condiția

$$\int_A Q(\omega, B)\, dP(\omega) = P(\tilde{\mathbf{x}}^{-1}(B) \cap A) \qquad (6.4.4)$$

pentru orice $A \in \mathcal{F}$.

Se poate arăta că, oricare ar fi vectorul aleator $\tilde{\mathbf{x}}$ pe $(E, \mathcal{K}, P)$ și oricare ar fi corpul borelian $\mathcal{F} \subset \mathcal{K}$, există o funcție Q care să fie repartiția regulată condiționată a vectorului aleator $\tilde{\mathbf{x}}$ relativ la corpul borelian $\mathcal{F}$ (vezi, de exemplu, Ghihman, I.I., Skorohod, A.V. (1971)).

Dacă Q este repartiția regulată condiționată a vectorului aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)'$ relativ la corpul borelian $\mathcal{F} \subset \mathcal{K}$, atunci funcția cu valori reale $Q(\omega; \tilde{x}_1 < x_1, \ldots, \tilde{x}_n < x_n)$ va fi numită *funcția de repartiție condiționată a vectorului* $\tilde{\mathbf{x}}$ *relativ la corpul borelian* $\mathcal{F}$ și va fi notată prin $F(x_1, \ldots, x_n \mid \mathcal{F})$. Se poate arăta că pentru orice funcție măsurabilă și integrabilă $g : R^n \to R$, avem

$$M(g(\tilde{x}_1, \ldots, \tilde{x}_n) \mid \mathcal{F}) = \int_{R^n} g(x_1, \ldots, x_n)\, dF(x_1, \ldots, x_n \mid \mathcal{F}). \qquad (6.4.5)$$

Scanned with OKEN Scanner

Să presupunem că $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)'$ este un vector aleator a cărui densitate de probabilitate este $f(x_1, \ldots, x_n)$. Dacă $m \leqslant n$, vom nota prin $\tilde{\mathbf{x}}_{(m)}$ vectorul cu componentele $\tilde{x}_1, \ldots, \tilde{x}_m$ și prin $f(x_1, \ldots, x_m)$ densitatea sa de probabilitate. Se știe că există relația

$$f(x_1, \ldots, x_m) = \int_{R^{n-m}} f(x_1, \ldots, x_n)\, \mathrm{d}x_{m+1} \ldots \mathrm{d}x_n. \tag{6.4.6}$$

În cele ce urmează ne propunem să calculăm valoarea medie condiționată $M(g(\tilde{x}_{m+1}, \ldots, \tilde{x}_n) \mid \tilde{x}_1, \ldots, \tilde{x}_m)$, unde $g : R^{n-m} \to R$ este măsurabilă, iar $g(\tilde{x}_{m+1}, \ldots, \tilde{x}_n)$ este integrabilă. Din definiția valorii medii condiționate rezultă că pentru orice $B \in \mathfrak{B}^m$ avem

$$\int_{\tilde{\mathbf{x}}_{(m)}^{-1}(B)} M(g(\tilde{x}_{m+1}, \ldots, \tilde{x}_n) \mid \tilde{x}_1, \ldots, \tilde{x}_m)\, \mathrm{d}P = \int_{\tilde{\mathbf{x}}_{(m)}^{-1}(B)} g(\tilde{x}_{m+1}, \ldots, \tilde{x}_n)\, \mathrm{d}P =$$

$$= M(g(\tilde{x}_{m+1}, \ldots, \tilde{x}_n)\chi_B(\tilde{x}_1, \ldots, \tilde{x}_m)). \tag{6.4.7}$$

Pe de altă parte, avem

$$M(g(\tilde{x}_{m+1}, \ldots, \tilde{x}_n)\chi_B(\tilde{x}_1, \ldots, \tilde{x}_m)) =$$

$$= \int_{R^n} g(x_{m+1}, \ldots, x_n)\chi_B(x_1, \ldots, x_m)\, \mathrm{d}x_1 \ldots \mathrm{d}x_n =$$

$$= \int_B \left(\int_{R^{n-m}} g(x_{m+1}, \ldots, x_n) \frac{f(x_1, \ldots, x_n)}{f(x_1, \ldots, x_m)}\, \mathrm{d}x_{m+1} \ldots \mathrm{d}x_n \right) \mathrm{d}x_1 \ldots \mathrm{d}x_m =$$

$$= \int_{\tilde{\mathbf{x}}_{(m)}^{-1}(B)} h(\tilde{x}_1, \ldots, \tilde{x}_m)\, \mathrm{d}P, \tag{6.4.8}$$

unde

$$h(x_1, \ldots, x_m) = \int_{R^{n-m}} g(x_{m+1}, \ldots, x_m) \frac{f(x_1, \ldots, x_n)}{f(x_1, \ldots, x_m)}\, \mathrm{d}x_{m+1} \ldots \mathrm{d}x_m.$$

Scanned with OKEN Scanner

Din (6.4.7) și (6.4.8) rezultă atunci

$$M(g(\tilde{x}_{m+1}, \ldots, \tilde{x}_n) \mid \tilde{x}_1, \ldots, \tilde{x}_m) =$$

$$= \int_{R^{n-m}} g(x_{m+1}, \ldots, x_m) \frac{f(\tilde{x}_1, \ldots, \tilde{x}_m, x_{m+1}, \ldots, x_n)}{f(\tilde{x}_1, \ldots, \tilde{x}_m)} \, dx_{m+1} \ldots dx_n.$$

Funcția

$$f(x_{m+1}, \ldots, x_n \mid x_1, \ldots, x_m) = \frac{f(x_1, \ldots, x_n)}{f(x_1, \ldots, x_m)} \tag{6.4.9}$$

care se pune în evidență în acest mod este numită *densitatea de probabilitate condiționată a vectorului aleator* $(x_{m+1}, \ldots, \tilde{x}_n)'$ *în ipoteza* $\tilde{x}_1 = x_1, \ldots, \tilde{x}_m = x_m$. Cu ajutorul densității de probabilitate condiționate valoarea medie condiționată se poate exprima sub forma

$$M(g(\tilde{x}_{m+1}, \ldots, \tilde{x}_n) \mid \tilde{x}_1, \ldots, \tilde{x}_m) =$$

$$= \int_{R^{n-m}} g(x_{m+1}, \ldots, x_n) f(x_{m+1}, \ldots, x_n \mid \tilde{x}_1, \ldots, \tilde{x}_m) \, dx_{m+1} \ldots dx_n. \tag{6.4.10}$$

Din cele de mai sus rezultă că densitatea de probabilitate condiționată $f(x_{m+1}, \ldots, x_n \mid x_1, \ldots, x_m)$ este definită pentru toate valorile posibile* $(x_1, \ldots, x_m)' \in R^m$ de relația (6.4.9), putînd lua valori arbitrare în caz contrar, de exemplu valoarea zero.

În aplicațiile practice se cunoaște în mod obișnuit $f(x_1, \ldots, x_m)$ și $f(x_{m+1}, \ldots, x_n \mid x_1, \ldots, x_m)$ și se cere determinarea densităților $f(x_1, \ldots, x_n)$ și $f(x_{m+1}, \ldots, x_n)$ și a densității condiționate $f(x_1, \ldots, x_m \mid x_{m+1}, \ldots, x_n)$. Este clar că $f(x_1, \ldots, x_n)$ rezultă din (6.4.9), punînd

$$f(x_1, \ldots, x_n) = \begin{cases} 0 & \text{dacă } f(x_1, \ldots, x_m) = 0, \\ f(x_1, \ldots, x_m) f(x_{m+1}, \ldots, x_n \mid x_1, \ldots, x_m) & \text{în caz contrar.} \end{cases} \tag{6.4.11}$$

---

* O valoare $(x_1, \ldots, x_m)' \in R^m$ a vectorului aleator $\tilde{x}_{(m)}$ este numită *posibilă* dacă $f(x_1, \ldots, x_m) > 0$.

Apoi, se determină $f(x_{m+1}, \ldots, x_n)$ pe calea obișnuită, punînd

$$f(x_{m+1}, \ldots, x_n) = \int_{R^m} f(x_1, \ldots, x_n)\, dx_1 \ldots dx_m \qquad (6.4.12)$$

și densitatea de probabilitate condiționată punînd

$$f(x_1, \ldots, x_m \mid x_{m+1}, \ldots, x_n) = \frac{f(x_1, \ldots, x_n)}{f(x_{m+1}, \ldots, x_n)} \qquad (6.4.13)$$

pentru $f(x_{m+1}, \ldots, x_n) > 0$ și zero în caz contrar.

Din formulele precedente rezultă

$$f(x_1, \ldots, x_m \mid x_{m+1}, \ldots, x_n) =$$

$$= \frac{f(x_1, \ldots, x_m) f(x_{m+1}, \ldots, x_n \mid x_1, \ldots, x_m)}{\int_{R^m} f(x_1, \ldots, x_m) f(x_{m+1}, \ldots, x_n \mid x_1, \ldots, x_m)\, dx_1 \ldots dx_m}, \qquad (6.4.14)$$

care reprezintă *formula lui Bayes pentru densitățile de probabilitate*. Această formulă ne permite să calculăm *densitatea de probabilitate aposteriori* $f(x_1, \ldots, x_m \mid x_{m+1}, \ldots, x_n)$ a variabilelor aleatoare $(\tilde{x}_1, \ldots, \tilde{x}_m)$ în funcție de *densitatea lor apriori* $f(x_1, \ldots, x_m)$ și de densitatea de probabilitate condiționată $f(x_{m+1}, \ldots, x_n \mid x_1, \ldots, x_m)$.

În sfîrșit, să observăm în încheiere că, integrînd (6.4.13) în raport cu variabilele $x_1, \ldots, x_m$, obținem

$$f(x_{m+1}, \ldots, x_n) = \int_{R^m} f(x_1, \ldots, x_m) f(x_{m+1}, \ldots, x_n \mid x_1, \ldots, x_m)\, dx_1 \ldots dx_m, \qquad (6.4.15)$$

de unde rezultă că funcția de repartiție a vectorului aleator $(\tilde{x}_{m+1}, \ldots, \tilde{x}_n)'$ este dată de relația

$$F(x_{m+1}, \ldots, x_n) = \int_{R^m} F(x_{m+1}, \ldots, x_n \mid x_1, \ldots, x_m) f(x_1, \ldots, x_m)\, dx_1 \ldots dx_m \qquad (6.4.16)$$

unde $F(x_{m+1}, \ldots, x_n \mid \tilde{x}_1, \ldots, \tilde{x}_m)$ este funcția de repartiție a vectorului aleator $(\tilde{x}_{m+1}, \ldots, \tilde{x}_n)'$ condiționată de $\tilde{x}_1, \ldots, \tilde{x}_m$.

Scanned with OKEN Scanner

6.5. MOMENTE CONDIȚIONATE

Să considerăm un cîmp borelian de probabilitate $(E,\mathcal{K},P)$, o variabilă aleatoare $\tilde{x}\in\mathcal{L}^r(E,\mathcal{K},P)$ și un corp borelian $\mathcal{F}\subset\mathcal{K}$ pe E.

Definiția 6.5.1. *Numim moment de ordinul* r *al variabilei aleatoare* $\tilde{x}$ *condiționat de* $\mathcal{F}$ *și îl notăm prin* $M_r(\tilde{x}\mid\mathcal{F})$ *media lui* $\tilde{x}^r$ *condiționată de* $\mathcal{F}$.

Avem deci prin definiție

$$M_r(\tilde{x}\mid\mathcal{F})=M(\tilde{x}^r\mid\mathcal{F}). \tag{6.5.1}$$

Definiția 6.5.2. *Numim moment centrat de ordinul* r *al variabilei aleatoare* $\tilde{x}$ *condiționat de* $\mathcal{F}$ *și îl notăm prin* $V_r(\tilde{x}\mid\mathcal{F})$ *media lui* $(\tilde{x}-M(\tilde{x}\mid\mathcal{F}))^r$ *condiționată de* $\mathcal{F}$.

Avem deci prin definiție

$$V_r(\tilde{x}\mid\mathcal{F})=M[(\tilde{x}-M(\tilde{x}\mid\mathcal{F}))^r\mid\mathcal{F}]. \tag{6.5.2}$$

Mai general, fie $\tilde{\mathbf{x}}=(\tilde{x}_1,\ldots,\tilde{x}_n)'$ un vector aleator pe cîmpul borelian de probabilitate $(E,\mathcal{K},P)$, $\mathcal{F}\subset\mathcal{K}$ un corp borelian pe E, $(r_1,\ldots,r_n)'\in R^n$ și $\tilde{x}_1^{r_1}\ldots\tilde{x}_n^{r_n}\in\mathcal{L}(E,\mathcal{K},P)$.

Definiția 6.5.3. *Numim moment de ordinul* $(r_1,\ldots,r_n)'$ *al vectorului aleator* $\tilde{\mathbf{x}}$ *condiționat de* $\mathcal{F}$ *și îl notăm prin* $M_{r_1\ldots r_n}(\tilde{\mathbf{x}}\mid\mathcal{F})$ *media lui* $\tilde{x}_1^{r_1}\ldots\tilde{x}_n^{r_n}$ *condiționată de* $\mathcal{F}$.

Avem deci prin definiție

$$M_{r_1\ldots r_n}(\tilde{\mathbf{x}}\mid\mathcal{F})=M(\tilde{x}_1^{r_1}\ldots\tilde{x}_n^{r_n}\mid\mathcal{F}). \tag{6.5.3}$$

Definiția 6.5.4. *Numim moment centrat de ordinul* $(r_1,\ldots,r_n)'$ *al vectorului aleator* $\tilde{\mathbf{x}}$ *condiționat de* $\mathcal{F}$ *și îl notăm prin* $V_{r_1\ldots r_n}(\tilde{\mathbf{x}}\mid\mathcal{F})$ *media variabilei aleatoare* $(\tilde{x}_1-M(\tilde{x}_1\mid\mathcal{F}))^{r_1}\ldots(\tilde{x}_n-M(\tilde{x}_n\mid\mathcal{F}))^{r_n}$ *condiționată de* $\mathcal{F}$.

Avem deci prin definiție

$$V_{r_1\ldots r_n}(\tilde{\mathbf{x}}\mid\mathcal{F})=M[((\tilde{x}_1-M(\tilde{x}_1\mid\mathcal{F}))^{r_1}\ldots(\tilde{x}_n-M(\tilde{x}_n\mid\mathcal{F}))^{r_n}\mid\mathcal{F}]. \tag{6.5.4}$$

Dintre momentele condiționate, cele mai importante sînt următoarele: media condiționată, dispersia condiționată și corelația (covarianța) condiționată. Dispersia lui $\tilde{x}$ condiționată de $\mathcal{F}$ este $V_2(\tilde{x}\mid\mathcal{F})$ și va mai fi notată și prin $V(\tilde{x}\mid\mathcal{F})$. Corelația variabilelor aleatoare $\tilde{x}$ și $\tilde{y}$ condiționată de $\mathcal{F}$ este prin definiție $V_{11}(\tilde{x},\tilde{y}\mid\mathcal{F})$, notată și prin $\operatorname{cov}(\tilde{x},\tilde{y}\mid\mathcal{F})$.

Scanned with OKEN Scanner

Se poate de asemenea defini coeficientul de corelație al variabilelor $\tilde{x}$ și $\tilde{y}$ condiționat de $\mathcal{F}$ prin formula

$$r(\tilde{x}, \tilde{y} \mid \mathcal{F}) = \frac{\operatorname{cov}(\tilde{x}, \tilde{y} \mid \mathcal{F})}{[V(\tilde{x} \mid \mathcal{F})\, V(\tilde{x} \mid \mathcal{F})]^{1/2}}. \tag{6.5.5}$$

Dispersia condiționată are proprietățile următoare:

$$V(\tilde{x}) = M(V(\tilde{x} \mid \mathcal{F})) + V(M(\tilde{x} \mid \mathcal{F})) \tag{6.5.6}$$

și

$$V(\tilde{x}) \geqslant M(V(\tilde{x} \mid \mathcal{F})). \tag{6.5.7}$$

Deoarece dispersia este nenegativă, proprietatea (6.5.7) decurge imediat din (6.5.6), ținînd seama că $V(M(\tilde{x} \mid \mathcal{F})) \geqslant 0$. Pentru a demonstra (6.5.6), să observăm mai întîi că putem scrie

$$V(\tilde{x}) = M[(\tilde{x} - M(\tilde{x} \mid \mathcal{F}))^2 + (M(\tilde{x} \mid \mathcal{F}) - M(\tilde{x}))^2] =$$

$$= M\{M[(\tilde{x} - M(\tilde{x} \mid \mathcal{F}))^2 + (M(\tilde{x} \mid \mathcal{F}) - M(\tilde{x}))^2 \mid \mathcal{F}]\}.$$

Dacă ridicăm la pătrat binomul și evaluăm media condiționată a fiecărui termen, conform definiției dispersiei condiționate, obținem

$$M((\tilde{x} - M(\tilde{x} \mid \mathcal{F}))^2 \mid \mathcal{F}) = V(\tilde{x} \mid \mathcal{F}),$$

și avînd în vedere proprietățile valorii medii condiționate

$$M[(M(\tilde{x} \mid \mathcal{F}) - M(\tilde{x}))^2 \mid \mathcal{F}] = [M(\tilde{x} \mid \mathcal{F}) - M(\tilde{x})]^2.$$

În plus, se observă că

$$M(\tilde{x}) = M(M(\tilde{x} \mid \mathcal{F}))$$

și

$$M[(\tilde{x} - M(\tilde{x} \mid \mathcal{F})(M(\tilde{x} \mid \mathcal{F}) - M(\tilde{x})) \mid \mathcal{F}] =$$

$$= [M(\tilde{x} \mid \mathcal{F}) - M(\tilde{x})]\, M[(\tilde{x} - M(\tilde{x} \mid \mathcal{F})) \mid \mathcal{F}] = 0,$$

de unde rezultă proprietatea (4.5.6).

Fie un vector aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)'$ pe $(E, \mathcal{K}, P)$, avînd componente integrabile și fie $\mathcal{F} \subset \mathcal{K}$ un corp borelian pe E.

Scanned with OKEN Scanner

Definiția 6.5.5. *Numim media lui* $\tilde{x}$ *condiționată de* $\mathcal{F}$ *și o notăm prin* $M(\tilde{x} \mid \mathcal{F})$ *vectorul cu componentele* $M(\tilde{x}_i \mid \mathcal{F})$, $1 \leqslant i \leqslant n$.

Definiția 6.5.6. *Numim matricea de covarianță a vectorului aleator* $\tilde{x}$ *condiționată de* $\mathcal{F}$ *și o notăm prin* $V(\tilde{x} \mid \mathcal{F})$ *matricea cu elementele* $\operatorname{cov}(\tilde{x}_i, \tilde{x}_j \mid \mathcal{F})$, $1 \leqslant i, j \leqslant n$.

Se poate arăta ușor că are loc relația

$$V(\tilde{x}) = M(V(\tilde{x} \mid \mathcal{F})) + V(M(\tilde{x} \mid \mathcal{F})), \tag{6.5.8}$$

proprietate care generalizează relația (6.5.6). Dacă se scrie relația (6.5.8) pe componente, se obține evident

$$\operatorname{cov}(\tilde{x}_i, \tilde{x}_j \mid \mathcal{F}) = M(\operatorname{cov}(\tilde{x}_i, \tilde{x}_j \mid \mathcal{F})) + \operatorname{cov}(M(\tilde{x}_i \mid \mathcal{F}), M(\tilde{x}_j \mid \mathcal{F})) \tag{6.5.9}$$

pentru $i \neq j$ și

$$V(\tilde{x}_i) = M(V(\tilde{x}_i \mid \mathcal{F})) + V(M(\tilde{x}_i \mid \mathcal{F})). \tag{6.5.10}$$

6.6. FUNCȚII DE REGRESIE

Să considerăm variabilele aleatoare $\tilde{x}_1, \ldots, \tilde{x}_n$ definite pe cîmpul de probabilitate $(E, \mathcal{K}, P)$.

Definiția 6.6.1. *Dacă* $\tilde{x}_i$, $1 \leqslant i \leqslant n$, *este integrabilă, atunci* $M(\tilde{x}_i \mid \tilde{x}_1, \ldots, \tilde{x}_{i-1}, \tilde{x}_{i+1}, \ldots, \tilde{x}_n)$, *privită ca funcție de* $\tilde{x}_1, \ldots, \tilde{x}_{i-1}, \tilde{x}_{i+1}, \ldots, \tilde{x}_n$, *este numită funcția de regresie a lui* $\tilde{x}_i$ *față de* $\tilde{x}_1, \ldots, \tilde{x}_{i-1}, \tilde{x}_{i+1}, \ldots, \tilde{x}_n$.

Regresia este *liniară* dacă $M(\tilde{x}_i \mid \tilde{x}_1, \ldots, \tilde{x}_{i-1}, \tilde{x}_{i+1}, \ldots, \tilde{x}_n)$ este funcție liniară de $\tilde{x}_1, \ldots, \tilde{x}_{i-1}, \tilde{x}_{i+1}, \ldots, \tilde{x}_n$, adică

$$M(\tilde{x}_i \mid \tilde{x}_1, \ldots, \tilde{x}_{i-1}, \tilde{x}_{i+1}, \ldots, \tilde{x}_n) = a_0 + \sum_{\substack{j=1 \\ j \neq i}}^{n} a_j \tilde{x}_j. \tag{6.6.1}$$

Coeficienții a_j, $0 \leqslant j \leqslant n$, $j \neq i$ care apar în (6.6.1) sînt numiți *coeficienți de regresie*.

Evident, dacă toate componentele $\tilde{x}_i$, $1 \leqslant i \leqslant n$, ale vectorului aleator $(\tilde{x}_1, \ldots, \tilde{x}_n)'$ sînt integrabile, se pot defini n funcții de regresie $M(\tilde{x}_i \mid \tilde{x}_1, \ldots, \tilde{x}_{i-1}, \tilde{x}_{i+1}, \ldots, \tilde{x}_n)$, $1 \leqslant i \leqslant n$.

Pentru vectorii aleatori cu numai două componente are loc rezultatul următor.

Scanned with OKEN Scanner

Teorema 6.6.1. *Dacă una dintre funcțiile de regresie ale vectorului aleator* $(\tilde{x}, \tilde{y})'$ *este constantă, atunci* $\tilde{x}$ *și* $\tilde{y}$ *sînt necorelate.*

Demonstrație. Să presupunem că $M(\tilde{x} \mid \tilde{y}) = k$, unde k este o constantă. Rezultă atunci că

$$M(\tilde{x}) = M(M(\tilde{x} \mid \tilde{y})) = k$$

și

$$M(\tilde{x}\tilde{y}) = M(M(\tilde{x}\tilde{y} \mid \tilde{y})) = M(\tilde{y}M(\tilde{x} \mid \tilde{y})) = k\,M(\tilde{y}).$$

Avem deci

$$\operatorname{cov}(\tilde{x}, \tilde{y}) = M(\tilde{x}\tilde{y}) - M(\tilde{x})\,M(\tilde{y}) = kM(\tilde{y}) - kM(\tilde{y}) = 0$$

și deci $\tilde{x}$ și $\tilde{y}$ sînt necorelate.

Teorema 6.6.2. *Dacă funcția de regresie a lui* $\tilde{x}_1$ *față de* $\tilde{x}_2$ *este liniară, adică*

$$M(\tilde{x}_1 \mid \tilde{x}_2) = a_0 + a_1\tilde{x}_2, \tag{6.6.2}$$

iar $V(\tilde{x}_2) \neq 0$, *atunci*

$$a_1 = r(\tilde{x}_1, \tilde{x}_2)[V(\tilde{x}_1)/V(\tilde{x}_2)]^{1/2} \tag{6.6.3}$$

și

$$a_0 = M(\tilde{x}_1) - M(\tilde{x}_2)\,r(\tilde{x}_1, \tilde{x}_2)\,[V(\tilde{x}_1)/V(\tilde{x}_2)]^{1/2}. \tag{6.6.4}$$

Demonstrație. Luînd valoarea medie în (6,6.2) obținem

$$M(\tilde{x}_1) = a_0 + a_1M(\tilde{x}_2). \tag{6.6.5}$$

Înmulțind (6.6.2) cu $\tilde{x}_2$ și luînd din nou valoarea medie obținem

$$M(\tilde{x}_1\tilde{x}_2) = a_0M(\tilde{x}_2) + a_1M(\tilde{x}_2^2). \tag{6.6.6}$$

Rezolvînd sistemul de ecuații (6.6.5) și (6.6.6) în raport cu a_0 și a_1 obținem rezultatele dorite.

Dacă notăm $f(\tilde{x}_1, \tilde{x}_2) = r$, $m_i = M(\tilde{x}_i)$, $\sigma_i^2 = V(\tilde{x}_i)$, $i = 1, 2$, atunci din (6.6.3) și (6.6.4) rezultă că

$$a_1 = r(\sigma_1/\sigma_2); \quad a_0 = m_1 - rm_2(\sigma_1/\sigma_2). \tag{6.6.7}$$

Evident, dacă regresia lui $\tilde{x}_2$ asupra lui $\tilde{x}_1$ este liniară, adică

$$M(\tilde{x}_2 \mid \tilde{x}_1) = a_0' + a_1'\tilde{x}_1,$$

și $V(\tilde{x}_1) \neq 0$, atunci rezultă ca mai sus că

$$a_1' = r(\sigma_2/\sigma_1); \quad a_0' = m_2 - rm_1(\sigma_2/\sigma_1). \tag{6.6.8}$$

Scanned with OKEN Scanner

Definiția 6.6.2. *Indiferent de tipul regresiei lui* $\tilde{x}_1$ *asupra lui* $\tilde{x}_2$, *dreapta de ecuație*

$$x_1 - m_1 = r(\sigma_1/\sigma_2)(x_2 - m_2) \tag{6.6.9}$$

este numită dreapta de regresie teoretică a lui $\tilde{x}_1$ *asupra lui* $\tilde{x}_2$.

Analog, dreapta de ecuație

$$x_2 - m_2 = r(\sigma_2/\sigma_1)(x_1 - m_1) \tag{6.6.10}$$

este numită *dreapta de regresie teoretică a lui* $\tilde{x}_2$ *asupra lui* $\tilde{x}_1$, indiferent dacă regresia lui $\tilde{x}_2$ asupra lui $\tilde{x}_1$ este liniară sau nu.

Considerațiile de mai sus se pot extinde ușor pentru un vector aleator $(\tilde{x}_1, \ldots, \tilde{x}_n)'$ cu $n \geqslant 2$ componente. Dacă funcția de regresie a lui $\tilde{x}_n$ față de $\tilde{x}_1, \ldots, \tilde{x}_{n-1}$ este liniară, adică

$$M(\tilde{x}_n \mid \tilde{x}_1, \ldots, \tilde{x}_{n-1}) = a_0 + a_1\tilde{x}_1 + \ldots + a_{n-1}\tilde{x}_{n-1}, \tag{6.6.11}$$

se pot de asemenea determina coeficienții de regresie a_i, $0 \leqslant i \leqslant n-1$.

Într-adevăr, notînd $M(\tilde{x}_i) = m_i$ și $\operatorname{cov}(\tilde{x}_i, \tilde{x}_j) = \sigma_{ij}$, $1 \leqslant i, j \leqslant n$, și trecînd la valoarea medie în (6.6.11) obținem

$$m_n = a_0 + a_1 m_1 + \ldots + a_{n-1} m_{n-1}. \tag{6.6.12}$$

Înlocuind în (6.6.11) pe a_0 cu valoarea care rezultă din (6.6.12) obținem

$$\begin{aligned} &M(\tilde{x}_n \mid \tilde{x}_1, \ldots, \tilde{x}_{n-1}) - m_n = \\ &= a_1(\tilde{x}_1 - m_1) + \ldots + a_{n-1}(\tilde{x}_{n-1} - m_{n-1}). \end{aligned} \tag{6.6.13}$$

Înmulțind ultima relație cu $\tilde{x}_i - m_i$ și trecînd la valoarea medie obținem

$$\sigma_{ni} = a_1\sigma_{1i} + \ldots + a_{n-1}\sigma_{n-1,i}, \quad 1 \leqslant i \leqslant n-1. \tag{6.6.14}$$

Presupunînd că vectorul aleator $(\tilde{x}_1, \ldots, \tilde{x}_{n-1})'$ este nedegenerat, rezultă că matricea $V(\tilde{x}_1, \ldots, \tilde{x}_{n-1}) = (\sigma_{ij})$, $1 \leqslant i, j \leqslant n-1$, este pozitiv definită și deci nesingulară. Rezolvînd sistemul de ecuații (6.6.14) obținem

$$a_j = (\det V(\tilde{x}_1, \ldots, \tilde{x}_{n-1}))^{-1} \sum_{i=1}^{n-1} \sigma_{ni} A_{ij}, \quad 1 \leqslant j \leqslant n-1, \tag{6.6.15}$$

unde A_{ij} este complementul algebric al elementului σ_{ji}.

Scanned with OKEN Scanner

Definiția 6.6.3. *Indiferent de tipul regresiei lui* $\tilde{x}_n$ *față de* $\tilde{x}_1, \ldots, \tilde{x}_{n-1}$, *vom numi hiperplan de regresie teoretic al variabilei* $\tilde{x}_n$ *asupra variabilelor* $\tilde{x}_1, \ldots, \tilde{x}_{n-1}$ *hiperplanul de ecuație*

$$x_n - m_n = (\det V(\tilde{x}_1, \ldots, \tilde{x}_{n-1}))^{-1} \sum_{j=1}^{n-1} \sum_{i=1}^{n-1} \sigma_{ni} A_{ij} (x_j - m_j). \quad (6.6.16)$$

Teorema 6.6.3. *Să considerăm funcția de* n *variabile reale definită în modul următor:*

$$g(c_0, \ldots, c_{n-1}) = M[(\tilde{x}_n - c_0 - c_1\tilde{x}_1 - \ldots - c_{n-1}\tilde{x}_{n-1})^2]. \quad (6.6.17)$$

Dacă vectorul aleator $(\tilde{x}_1, \ldots, \tilde{x}_{n-1})'$ *este nedegenerat, atunci minimul funcției* g *este atins pentru* $c_i = a_i$, $0 \leqslant i \leqslant n-1$, *unde* a_i, $0 \leqslant i \leqslant n-1$, *sînt coeficienții de regresie liniară a lui* $\tilde{x}_n$ *față de* $\tilde{x}_1, \ldots, \tilde{x}_{n-1}$. *Valoarea minimă a lui* g, *pe care o vom nota prin* $\sigma_{n/12\ldots n-1}$, *este dată de relația*

$$\sigma_{n/12\ldots n-1} = (\det V(\tilde{x}_1, \ldots, \tilde{x}_n))/(\det V(\tilde{x}_1, \ldots, \tilde{x}_{n-1})). \quad (6.6.18)$$

Demonstrație. Punctul de minim al lui g este soluția sistemului de ecuații

$$\frac{\partial g(c_0, \ldots, c_{n-1})}{\partial c_i} = 0, \quad 0 \leqslant i \leqslant n-1. \quad (6.6.19)$$

Pentru $i = 0$ se obține

$$m_n = c_0 + c_1 m_1 + \ldots + c_{n-1} m_{n-1}. \quad (6.6.20)$$

Dacă se înlocuiește valoarea lui c_0 dedusă din (6.6.20) în celelalte $n-1$ ecuații ale sistemului (6.6.19), se obține

$$\sigma_{ni} = c_1\sigma_{1i} + \ldots + c_{n-1}\sigma_{n-1,\,i}, \quad 1 \leqslant i \leqslant n-1,$$

a cărui soluție este chiar $c_i = a_i$, $1 \leqslant i \leqslant n-1$, ceea ce demonstrează prima afirmație a teoremei.

Valoarea minimă a lui g este deci

$$g(a_0, \ldots, a_{n-1}) = M\left\{\left[\sum_{i=1}^{n-1} a_i(\tilde{x}_i - m_i) - (\tilde{x}_n - m_n)\right]^2\right\} =$$

$$= \sigma_n^2 - 2\sum_{j=1}^{n-1} a_j\sigma_{nj} + \sum_{i=1}^{n-1}\sum_{j=1}^{n-1} a_i a_j \sigma_{ij}. \quad (6.6.21)$$

Scanned with OKEN Scanner

Ținînd seama de valorile coeficienților de regresie a_j, $1 \leqslant j \leqslant n-1$, rezultă că

$$\sum_{j=1}^{n-1} a_j \sigma_{nj} = (\det V(\tilde{x}_1, \ldots, \tilde{x}_{n-1}))^{-1} \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \sigma_{ni} \sigma_{nj} A_{ij} \quad (6.6.22)$$

și

$$\sum_{i=1}^{n-1} \sum_{j=1}^{n-1} a_i a_j \sigma_{ij} = (\det V(\tilde{x}_1, \ldots, \tilde{x}_{n-1}))^{-1} \sum_{i=1}^{-1} \sum_{j=1}^{n-1} \sum_{r=1}^{n-1} \sum_{s=1}^{n-1} \sigma_{ij} \sigma_{nr} \sigma_{ns} A_{ri} A_{sj}. \quad (6.6.23)$$

Deoarece

$$\sum_{j=1}^{n-1} \sigma_{ij} A_{sj} = \delta_{is} \det V(\tilde{x}_1, \ldots, \tilde{x}_{n-1}), \quad (6.6.24)$$

unde δ_{is} este simbolul lui Krönecker, rezultă că expresiile (6.6.22) și (6.6.23) sînt egale și deci (6.6.21) devine

$$g(a_0, \ldots, a_{n-1}) = \frac{\sigma_n^2 \det V(\tilde{x}_1, \ldots, \tilde{x}_{n-1}) - \sum_{i=1}^{n-1} \sigma_{ni} \sum_{j=1}^{n-1} \sigma_{nj} A_{ij}}{\det V(\tilde{x}_1, \ldots, \tilde{x}_{n-1})} =$$

$$= \frac{\det V(\tilde{x}_1, \ldots, \tilde{x}_n)}{\det V(\tilde{x}_1, \ldots, \tilde{x}_{n-1})}, \quad (6.6.25)$$

dacă se ține seama că

$$-\sum_{j=1}^{n-1} \sigma_{nj} A_{ij}$$

reprezintă complementul algebric al lui σ_{ni} în $\det V(\tilde{x}_1, \ldots, \tilde{x}_n)$. În sfîrșit, din (6.6.25) rezultă (6.6.18) și teorema este complet demonstrată.

În cazul particular $n = 2$ obținem

$$\frac{\det V(\tilde{x}_1, \tilde{x}_2)}{V(\tilde{x}_1)} = \sigma_2^2 (1 - r^2).$$

În încheiere vom prezenta noțiunile de *raport de corelație multiplă* și *coeficient de corelație multiplă*.

Numim *raport de corelație multiplă* a lui $\tilde{x}_n$ față de $\tilde{x}_1, \ldots, \tilde{x}_{n-1}$ și îl notăm prin $\eta^2_{n/12\ldots n-1}$ raportul

$$\eta^2_{n/12\ldots n-1} = \frac{M(V(\tilde{x}_n \mid \tilde{x}_1, \ldots, \tilde{x}_{n-1}))}{V(\tilde{x}_n)}. \tag{6.6.26}$$

Dintr-un rezultat cunoscut (vezi relația (6.5.7)) reiese că raportul de corelație multiplă este mai mic sau egal cu 1; evident, el este nenegativ. Se vede de asemenea ușor că $\eta^2_{n/12\ldots n-1} = 1$ dacă și numai dacă $\tilde{x}_n$ este independentă de $\tilde{x}_1, \ldots, \tilde{x}_{n-1}$. În cazul particular $n = 2$, numărul $\eta^2_{2/1}$ este numit *raport de corelație*.

Coeficientul de corelație dintre variabila aleatoare $\tilde{x}_n$ și variabila aleatoare

$$m_n + \sum_{j=1}^{n} a_j(\tilde{x}_j - m_j)$$

corespunzătoare regresiei liniare a lui $\tilde{x}_n$ față de $\tilde{x}_1, \ldots, \tilde{x}_{n-1}$ este numit *coeficient de corelație multiplă* și este notat prin $r^2_{n/12\ldots n-1}$. Avem deci prin definiție

$$r^2_{n/12\ldots n-1} = \frac{\operatorname{cov}\left(\tilde{x}_n,\ m_n + \sum_{j=1}^{n-1} a_j(\tilde{x}_j - m_j)\right)}{\left[V(\tilde{x}_n)V\left(m_n + \sum_{j=1}^{n-1} a_j(\tilde{x}_j - m_j)\right)\right]^{1/2}}. \tag{6.6.27}$$

Evident, coeficientul de corelație multiplă este cuprins între -1 și $+1$. Deoarece

$$V(\tilde{x}_n) = \sigma_{nn} = \sigma_n^2,$$

$$V\left(m_n + \sum_{j=1}^{n-1} a_j(\tilde{x}_j - m_j)\right) = \sum_{i=1}^{n-1}\sum_{j=1}^{n-1} \sigma_{ij}a_i a_j \tag{6.6.28}$$

și

$$\operatorname{cov}\left(\tilde{x}_n,\ m_n + \sum_{j=1}^{n-1} a_j(\tilde{x}_j - m_j)\right) = \sum_{j=1}^{n-1} \sigma_{nj}a_j, \tag{6.6.29}$$

rezultă că

$$r^2_{n/12\ldots n-1} = \frac{\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\sigma_{ni}\sigma_{nj}A_{ij}}{\sigma_n^2 \det V(\tilde{x}_1, \ldots, \tilde{x}_{n-1})}, \qquad (6.6.30)$$

dacă se utilizează faptul că (6.6.28) și (6.6.29) au valori egale, conform unei observații anterioare.

Deoarece

$$\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\sigma_{ni}\sigma_{nj}A_{ij} = \sigma_n^2 \det V(\tilde{x}_1, \ldots, \tilde{x}_{n-1}) - \det V(\tilde{x}_1, \ldots, \tilde{x}_n),$$

din (6.6.30) rezultă

$$r^2_{n/12\ldots n-1} = 1 - \frac{\det V(\tilde{x}_1, \ldots, \tilde{x}_n)}{\det V(\tilde{x}_1, \ldots, \tilde{x}_{n-1})} \qquad (6.6.31)$$

care este formula căutată.

Capitolul 7

ȘIRURI DE VARIABILE ALEATOARE

Pentru șirurile de variabile aleatoare pot fi introduse mai multe tipuri de convergență. În cele ce urmează vom prezenta succint aceste tipuri de convergență, precum și legăturile dintre ele.

7.1. CONVERGENȚA APROAPE SIGURĂ

Fie un șir de variabile aleatoare $(\tilde{x}_n)$ pe cîmpul borelian de probabilitate $(E,\mathcal{K},P)$ și o variabilă aleatoare $\tilde{x}$ pe același cîmp de probabilitate.

Definiția 7.1.1. *Spunem că șirul* $(\tilde{x}_n)$ *converge aproape sigur către* $\tilde{x}$ *și notăm* $\tilde{x}_n \xrightarrow{a.s.} \tilde{x}$, *dacă mulțimea acelor* $\omega \in E$ *pentru care șirul de numere reale* $(\tilde{x}_n(\omega))$ *converge spre numărul real* $\tilde{x}(\omega)$ *este de probabilitate* 1.

Cu alte cuvinte, $\tilde{x}_n \xrightarrow{a.s.} \tilde{x}$ dacă

$$P(\{\omega : \tilde{x}_n(\omega) \to \tilde{x}(\omega)\}) = 1.$$

Evident, rezultă că

$$P(\{\omega : \tilde{x}_n(\omega) \not\to \tilde{x}(\omega)\}) = 0.$$

Se constată ușor că dacă $\tilde{x}_n \xrightarrow{a.s.} \tilde{x}$ și $\tilde{y}_n \xrightarrow{a.s.} \tilde{y}$, atunci $\tilde{x} = \tilde{y}$ aproape sigur. Reciproc, dacă $\tilde{x}_n \xrightarrow{a.s.} \tilde{x}$ și $\tilde{x} = \tilde{y}$ aproape sigur, atunci $\tilde{x}_n \xrightarrow{a.s.} \tilde{y}$.

Dacă notăm prin $A \subset E$ mulțimea punctelor pentru care șirul $\tilde{x}_n(\omega)$ converge spre $\tilde{x}(\omega)$, adică

$$A = \{\omega : \tilde{x}_n(\omega) \to \tilde{x}(\omega)\}, \qquad (7.1.1)$$

putem găsi o formă utilă a acestei mulțimi. Fixîndu-ne la cazul variabilelor aleatoare finite*, putem arăta că A este de forma

$$A = \bigcap_{\varepsilon>0} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \{\omega : |\tilde{x}_m(\omega) - \tilde{x}(\omega)| < \varepsilon\}. \qquad (7.1.2)$$

Într-adevăr, dacă $\omega \in A$, rezultă că $\tilde{x}_n(\omega)$ converge spre $\tilde{x}(\omega)$ și deci pentru orice $\varepsilon > 0$ putem determina $N(\varepsilon, \omega)$ astfel încît să avem

$$|\tilde{x}_n(\omega) - \tilde{x}(\omega)| < \varepsilon$$

pentru orice $n \geqslant N(\varepsilon, \omega)$. Cu alte cuvinte, avem

$$\omega \in \bigcap_{m=N(\varepsilon,\,\omega)}^{\infty} \{\omega : |\tilde{x}_m(\omega) - \tilde{x}(\omega)| < \varepsilon\}$$

pentru orice $\varepsilon > 0$ și deci

$$\omega \in \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \{\omega : |\tilde{x}_m(\omega) - \tilde{x}(\omega)| < \varepsilon\}.$$

pentru orice $\varepsilon > 0$, de unde rezultă că

$$A \subset \bigcap_{\varepsilon>0} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \{\omega : |\tilde{x}_m(\omega) - \tilde{x}(\omega)| < \varepsilon\}$$

Incluziunea inversă se obține imediat.

Definiția 7.1.2. *Spunem că șirul de variabile aleatoare* $(\tilde{x}_n)$ *converge tare spre variabila aleatoare* $\tilde{x}$, *dacă*

$$\lim_n P\left(\bigcup_{m=0}^{\infty} \{\omega : |\tilde{x}_{n+m}(\omega) - \tilde{x}(\omega)| \geqslant \varepsilon\}\right) = 0 \qquad (7.1.3)$$

pentru orice $\varepsilon > 0$.

Este ușor de văzut că acest tip de convergență este echivalent cu convergența aproape sigură. Într-adevăr, dacă $\tilde{x}_n \xrightarrow{a.s.} \tilde{x}$, rezultă că mulțimea A de convergență este dată de (7.1.2) și deci avem

$$A \subset \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \{\omega : |\tilde{x}_m(\omega) - \tilde{x}(\omega)| < \varepsilon\}.$$

---

* Cazul variabilelor aleatoare oarecare se tratează similar.

Scanned with OKEN Scanner

Rezultă atunci că

$$P(A) \leqslant \lim_n P\left(\bigcap_{m=0}^{\infty} \{\omega : |\tilde{x}_{n+m}(\omega) - \tilde{x}(\omega)| < \varepsilon\}\right)$$

și deci

$$1 - P(A) \geqslant 1 - \lim_n P\left(\bigcup_{m=0}^{\infty} \{\omega : |\tilde{x}_{n+m}(\omega) - \tilde{x}(\omega)| \geqslant \varepsilon\}\right). \qquad (7.1.4)$$

Deoarece $P(A) = 1$, din inegalitatea (7.1.4) se deduce imediat relația (7.1.3), ceea ce arată că $(\tilde{x}_n)$ converge tare spre $\tilde{x}$.

Reciproc, presupunînd că $(\tilde{x}_n)$ converge tare spre $\tilde{x}$, obținem

$$P(A) = P\left(\bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \{\omega : |\tilde{x}_m(\omega) - \tilde{x}(\omega)| < 1/k\}\right) =$$

$$= \lim_k P\left(\bigcup_{n=1}^{\infty} \bigcap_{m=0}^{\infty} \{\omega : |\tilde{x}_{n+m}(\omega) - \tilde{x}(\omega)| < 1/k\}\right) =$$

$$= \lim_k \lim_n P\left(\bigcap_{m=0}^{\infty} \{\omega : |\tilde{x}_{n+m}(\omega) - \tilde{x}(\omega)| < 1/k\}\right) = 1$$

și deci $\tilde{x}_n \xrightarrow{\text{a.s.}} \tilde{x}$.

7.2. CONVERGENȚA ÎN PROBABILITATE

Fie $(\tilde{x}_n)$ un șir de variabile aleatoare pe cîmpul borelian de probabilitate $(E, \mathcal{K}, P)$ și fie $\tilde{x}$ o variabilă aleatoare pe același cîmp de probabilitate.

Definiția 7.2.1. *Spunem că șirul* $(\tilde{x}_n)$ *converge în probabilitate către* $\tilde{x}$ *și notăm* $\tilde{x}_n \xrightarrow{P} \tilde{x}$ *sau* $P\text{-}\lim_n \tilde{x}_n = \tilde{x}$, *dacă*

$$\lim_n P(\{\omega : |\tilde{x}_n(\omega) - \tilde{x}(\omega)| \geqslant \varepsilon\}) = 0 \qquad (7.2.1)$$

pentru orice $\varepsilon > 0$.

Scanned with OKEN Scanner

Cu alte cuvinte, pentru orice $\varepsilon > 0$ și $\delta > 0$ există un număr natural $N(\varepsilon, \delta)$ astfel încît

$$P(\{\omega : | \tilde{x}_n(\omega) - \tilde{x}(\omega) | \geqslant \varepsilon\}) < \delta \qquad (7.2.2)$$

pentru orice $n \geqslant N(\varepsilon, \delta)$.

Se constată ușor că dacă $\tilde{x}_n \xrightarrow{p} \tilde{x}$ și $\tilde{x}_n \xrightarrow{p} \tilde{y}$, atunci $\tilde{x} = \tilde{y}$ aproape sigur. Într-adevăr, deoarece

$$\{\omega : | \tilde{x}(\omega) - \tilde{y}(\omega) | \geqslant \varepsilon\} \subset \{\omega : | \tilde{x}_n(\omega) - \tilde{x}(\omega) \geqslant \varepsilon/2\} \cup \{\omega : | \tilde{x}_n(\omega) - \tilde{y}(\omega) | \geqslant \varepsilon/2\},$$

rezultă

$$P(|\tilde{x} - \tilde{y}| \geqslant \varepsilon) \leqslant P(|\tilde{x}_n - \tilde{x}| \geqslant \varepsilon/2) + P(|\tilde{x}_n - \tilde{y}| \leqslant \varepsilon/2),$$

de unde trecînd la limită obținem

$$P(|\tilde{x} - \tilde{y}| \geqslant \varepsilon) = 0$$

pentru orice $\varepsilon > 0$. Deoarece

$$P(\tilde{x} \neq \tilde{y}) = P\left(\bigcup_{n=1}^{\infty} \{\omega : | \tilde{x}(\omega) - \tilde{y}(\omega) | \geqslant 1/n\}\right) \leqslant \sum_{n=1}^{\infty} P(|\tilde{x} - \tilde{y}| \geqslant 1/n)$$

rezultă $\tilde{x} = \tilde{y}$ aproape sigur.

Analog se poate arăta că dacă $\tilde{x}_n \xrightarrow{p} \tilde{x}$ și $\tilde{x} = \tilde{y}$ aproape sigur, atunci $\tilde{x}_n \xrightarrow{p} \tilde{y}$.

Legătura dintre convergența aproape sigură și convergența în probabilitate este precizată în teorema următoare.

Teorema 7.2.1. *Fie* $(\tilde{x}_n)$ *un șir de variabile aleatoare pe cîmpul borelian de probabilitate* $(E, \mathfrak{K}, P)$ *și fie* $\tilde{x}$ *o variabilă aleatoare pe același cîmp de probabilitate.*

(a) Dacă $\tilde{x}_n \xrightarrow{a.s.} \tilde{x}$, *atunci* $\tilde{x}_n \xrightarrow{p} \tilde{x}$;

(b) (F. Riesz) *Dacă* $\tilde{x}_n \xrightarrow{p} \tilde{x}$, *atunci există un subșir* $(\tilde{x}_{n_k})$ *al șirului* $(\tilde{x}_n)$ *care converge aproape sigur către* $\tilde{x}$.

Demonstrație. (a) Dacă $\tilde{x}_n \xrightarrow{a.s.} \tilde{x}$, atunci rezultă că $(\tilde{x}_n)$ converge tare spre $\tilde{x}$, adică

$$\lim_n P\left(\bigcup_{m=0}^{\infty} \{\omega : | \tilde{x}_{n+m}(\omega) - \tilde{x}(\omega) | \geqslant \varepsilon\}\right) = 0$$

pentru orice $\varepsilon > 0$, de unde rezultă imediat că

$$\lim_n P(\{\omega : | \tilde{x}_n(\omega) - \tilde{x}(\omega) | \geqslant \varepsilon) = 0$$

pentru orice $\varepsilon > 0$, ceea ce arată că $\tilde{x}_n \xrightarrow{p} \tilde{x}$. Evident, demonstrația dată aici este valabilă numai pentru variabile aleatoare finite. Rezultatul este însă adevărat în general.

(*b*) Deoarece $\tilde{x}_n \xrightarrow{p} \tilde{x}$, rezultă că pentru fiecare întreg nenegativ k se poate determina un număr natural n_k cu proprietatea ca probabilitatea evenimentului

$$A_k = \{\omega: | \tilde{x}_{n_k}(\omega) - \tilde{x}(\omega) | \geqslant 1/(k+1)\}$$

să fie mai mică decît $2^{-(k+1)}$. Se poate vedea ușor că putem alege $n_k < n_{k+1}$ pentru $k = 0, 1, \ldots$

Evenimentele

$$B_j = \bigcup_{k=j}^{\infty} A_k, \quad j = 0, 1, 2, \ldots$$

formează un șir descrescător de evenimente și deci

$$P\left(\bigcap_{j=0}^{\infty} B_j\right) = \lim_j P(B_j) \leqslant \lim_j \sum_{k=j}^{\infty} P(A_k) \leqslant \lim_j \sum_{k=j}^{\infty} 2^{-(k+1)} = 0,$$

ceea ce arată că evenimentul

$$C = \bigcup_{j=0}^{\infty} \bigcap_{k=j}^{\infty} \{\omega : | \tilde{x}_{n_k}(\omega) - \tilde{x}(\omega) | < 1/(k+1)\}$$

are probabilitatea 1.

Evident, dacă $\omega \in C$, atunci șirul de numere reale $\tilde{x}_n(\omega)$ converge spre $\tilde{x}(\omega)$. Într-adevăr, din faptul că $\omega \in C$ rezultă că există $k_0 \geqslant 0$ astfel încît

$$| \tilde{x}_{n_k}(\omega) - \tilde{x}(\omega) | < 1/(k+1)$$

pentru orice $k \geqslant k_0$. Dacă $\varepsilon > 0$ este dat, atunci

$$| \tilde{x}_{n_k}(\omega) - \tilde{x}(\omega) | < \varepsilon$$

pentru orice $k \geqslant k_1(\varepsilon)$, unde $k_1(\varepsilon) = \max(k_0, 1/\varepsilon)$, ceea ce încheie demonstrația teoremei.

Scanned with OKEN Scanner

Teorema 7.2.2. *Șirul de variabile aleatoare* $(\widetilde{x}_n)$ *converge în probabilitate spre* $\widetilde{x}$ *dacă și numai dacă*

$$\lim_n M\left(\frac{|\widetilde{x}_n - \widetilde{x}|}{1 + |\widetilde{x}_n - \widetilde{x}|}\right) = 0. \qquad (7.2.3)$$

Demonstrație. Să considerăm funcția de variabilă reală

$$g(x) = \frac{|x|}{1 + |x|}.$$

Se vede ușor că g este nenegativă, pară, restricția sa la $[0, +\infty)$ este crescătoare, iar $\widetilde{y} = g(\widetilde{x})$ este o variabilă aleatoare integrabilă, oricare ar fi variabila aleatoare $\widetilde{x}$. Aplicînd inegalitatea (5.4.12) obținem

$$M\left(\frac{|\widetilde{x}_n - \widetilde{x}|}{1 + |\widetilde{x}_n - \widetilde{x}|}\right) -$$

$$-\frac{\varepsilon}{1+\varepsilon} \leqslant P(\{\omega : |\widetilde{x}_n - \widetilde{x}| \geqslant \varepsilon) \leqslant \frac{1+\varepsilon}{\varepsilon} M\left(\frac{|\widetilde{x}_n - \widetilde{x}|}{1 + |\widetilde{x}_n - \widetilde{x}|}\right),$$

de unde rezultă imediat teorema.

Observația 1. Dacă pentru două variabile aleatoare $\widetilde{x}$ și $\widetilde{y}$ pe $(E, \mathfrak{K}, P)$, punem

$$\rho(\widetilde{x}, \widetilde{y}) = M\left(\frac{|\widetilde{x} - \widetilde{y}|}{1 + |\widetilde{x} - \widetilde{y}|}\right), \qquad (7.2.4)$$

se poate vedea ușor că ρ este o semimetrică pe mulțimea variabilelor aleatoare pe $(E, \mathfrak{K}, P)$. Din teorema (7.2.2) rezultă atunci că convergența în probabilitate este echivalentă cu convergența în semimetrica ρ.

Observația 2. Din inegalitatea lui Cebîșev rezultă că

$$P(\{\omega : |\widetilde{x}_n(\omega) - \widetilde{x}(\omega)| \geqslant \varepsilon\}) \leqslant \frac{M_2(\widetilde{x}_n - \widetilde{x})}{\varepsilon^2} \qquad (7.2.5)$$

și deci convergența spre zero a momentului de ordinul al doilea al variabilei aleatoare $(\widetilde{x}_n - \widetilde{x})$ este o condiție suficientă ca $(\widetilde{x}_n)$ să conveargă spre $\widetilde{x}$ în probabilitate.

Scanned with OKEN Scanner

Teorema 7.2.3. *Fie* $(\tilde{x}_n)$ *și* $(\tilde{y}_n)$ *șiruri de variabile aleatoare convergente în probabilitate spre variabilele aleatoare* $\tilde{x}$ *și respectiv* $\tilde{y}$, *iar* a *și* b *numere reale. În aceste condiții rezultă că*

$$\text{P-}\lim_n |\tilde{x}_n| = |\tilde{x}| \tag{7.2.6}$$

și

$$\text{P-}\lim_n (a\tilde{x}_n + b\tilde{y}_n) = a\tilde{x} + b\tilde{y}. \tag{7.2.7}$$

Demonstrație. Pentru demonstrarea primei afirmații, să observăm că pentru orice $\varepsilon > 0$ avem

$$\{\omega : ||\tilde{x}_n(\omega)| - |\tilde{x}(\omega)|| \geqslant \varepsilon\} \subset \{\omega : |\tilde{x}_n(\omega) - \tilde{x}(\omega)| \geqslant \varepsilon\}$$

și deci

$$P(\{\omega : ||\tilde{x}_n(\omega)| - |\tilde{x}(\omega)|| \geqslant \varepsilon\}) \leqslant P(\{\omega : |\tilde{x}_n(\omega) - \tilde{x}(\omega)| \geqslant \varepsilon\}),$$

de unde rezultă imediat (7.2.6).

Pentru demonstrarea celei de a doua afirmații, să observăm că

$$\{\omega : |a\tilde{x}_n(\omega) + b\tilde{y}_n(\omega) - a\tilde{x}(\omega) - b\tilde{y}(\omega)| \geqslant \varepsilon\} \subset \{\omega : |a| \cdot |\tilde{x}_n(\omega) - \tilde{x}(\omega)| \geqslant \varepsilon/2\} \cup \{\omega : |b| \cdot |\tilde{y}_n(\omega) - \tilde{y}(\omega)| \geqslant \varepsilon/2\}$$

de unde rezultă

$$P(|a\tilde{x}_n + b\tilde{y}_n - a\tilde{x} - b\tilde{y}| \geqslant \varepsilon) \leqslant P(|a| \cdot |\tilde{x}_n - \tilde{x}| \geqslant \varepsilon/2) + P(|b| \cdot |\tilde{y}_n - \tilde{y}| \geqslant \varepsilon/2). \tag{7.2.8}$$

Dacă $a = 0$ sau $b = 0$, rezultatul (7.2.7) se obține imediat. Dacă $ab \neq 0$, atunci (7.2.7) se obține ușor din (7.2.8).

Observație. Se poate demonstra un rezultat mult mai general. Mai precis, fie $(\tilde{x}_n^i)$, $1 \leqslant i \leqslant m$, m șiruri de variabile aleatoare convergente în probabilitate către variabilele aleatoare $\tilde{x}^i$, $1 \leqslant i \leqslant m$. Dacă $g : R^m \to R$ este o funcție continuă pe R^m, cu excepția eventual a punctelor unei mulțimi $B \in \mathfrak{B}^m$ pentru care

$$P((\tilde{x}^1, \ldots, \tilde{x}^m) \in B) = 0,$$

atunci șirul de variabile aleatoare $(g(\tilde{x}_n^1, \ldots, \tilde{x}_n^m))$ converge în probabilitate spre $g(\tilde{x}^1, \ldots, \tilde{x}^m)$, adică

$$\text{P-}\lim_n g(\tilde{x}_n^1, \ldots, \tilde{x}_n^m) = g(\tilde{x}^1, \ldots, \tilde{x}^m). \tag{7.2.9}$$

Scanned with OKEN Scanner

În particular, rezultă că, în ipotezele teoremei (7.2.3), șirul $(\widetilde{x}_n\widetilde{y}_n)$ converge în probabilitate spre $(\widetilde{x}\widetilde{y})$, iar $(\widetilde{x}_n/\widetilde{y}_n)$ converge în probabilitate spre $(\widetilde{x}/\widetilde{y})$, ultima afirmație avînd loc cu presupunerea suplimentară că $P(P\text{-}\lim_n \widetilde{y}_n = 0) = 0$.

7.3. CONVERGENȚA ÎN MEDIE

Fie $(E,\mathcal{K},P)$ un cîmp borelian de probabilitate și fie $\mathcal{L}^r(E,\mathcal{K},P)$, $0<r<+\infty$, mulțimea variabilelor aleatoare $\widetilde{x}$ definite pe $(E, \mathcal{K}, P)$ care au proprietatea că $|\widetilde{x}|^r$ este integrabilă, adică $M(|\widetilde{x}|^r) < +\infty$. Pentru $\widetilde{x}, \widetilde{y}\in\mathcal{L}^r(E, \mathcal{K}, P)$ definim

$$\rho_r(\widetilde{x},\widetilde{y}) = [M(|\widetilde{x} - \widetilde{y}|^r)]^{1/r}. \tag{7.3.1}$$

Din proprietățile momentelor (vezi § 5.4) rezultă că ρ_r este o semimetrică pe $\mathcal{L}^r(E,\mathcal{K},P)$. Într-adevăr, este evident că $\rho_r(\widetilde{x}, \widetilde{y}) \geqslant 0$ și $\rho_r(\widetilde{x}, \widetilde{x}) = 0$ pentru orice $\widetilde{x}, \widetilde{y}\in\mathcal{L}^r(E,\mathcal{K},P)$. Din inegalitatea lui Minkowski (5.4.10) obținem

$$\rho_r(\widetilde{x}, \widetilde{y})=[\overline{M}_r(\widetilde{x}-\widetilde{y})]^{1/r}=[\overline{M}_r((\widetilde{x}-\widetilde{z})+(\widetilde{z}-\widetilde{y}))]^{1/r}\leqslant[\overline{M}_r(\widetilde{x}-\widetilde{z})]^{1/r}+$$
$$+[\overline{M}_r(\widetilde{z}-\widetilde{y})]^{1/r} = \rho_r(\widetilde{x}, \widetilde{z}) + \rho_r(\widetilde{y}, \widetilde{z})$$

pentru orice $\widetilde{x}, \widetilde{y}, \widetilde{z}\in\mathcal{L}^r(E,\mathcal{K},P)$, ceea ce încheie demonstrația.

Definiția 7.3.1. *Spunem că șirul de variabile aleatoare* $(\widetilde{x}_n)\subset\mathcal{L}^r(E, \mathcal{K}, P)$ *converge în medie de ordinul* r *spre* $\widetilde{x}\in\mathcal{L}^r(E, \mathcal{K}, P)$ *dacă*

$$\lim_n \rho_r(\widetilde{x}_n, \widetilde{x}) = 0. \tag{7.3.2}$$

Din definiție rezultă că $(\widetilde{x}_n)$ converge în medie de ordinul r spre $\widetilde{x}$ dacă și numai dacă

$$\lim_n M(|\widetilde{x}_n - \widetilde{x}|^r) = 0. \tag{7.3.3}$$

Din proprietatea (5.4.10) a momentelor rezultă că dacă $(\widetilde{x}_n)$ converge în medie de ordinul r spre $\widetilde{x}$ și dacă $0 < r' \leqslant r$, atunci $(\widetilde{x}_n)$ converge în medie de ordinul r' spre $\widetilde{x}$.

Scanned with OKEN Scanner

Teorema 7.3.1. *Dacă șirul* $(\tilde{x}_n) \subset \mathfrak{L}^r(E, \mathcal{K}, P)$ *converge în medie de ordinul* r *spre* $\tilde{x} \in \mathfrak{L}^r(E, \mathcal{K}, P)$, *atunci* $(\tilde{x}_n)$ *converge în probabilitate spre* $\tilde{x}$.

Demonstrație. Aplicînd inegalitatea lui Markov (vezi § 5.4) șirului de variabile aleatoare $(\tilde{x}_n - \tilde{x})$ obținem

$$P(\{\omega : |\tilde{x}_n(\omega) - \tilde{x}(\omega)| \geqslant \varepsilon\}) \leqslant \frac{M(|\tilde{x}_n - \tilde{x}|^r)}{\varepsilon^r} \qquad (7.3.4)$$

pentru orice $\varepsilon > 0$. Trecînd la limită în (7.3.4) și ținînd seama de (7.3.3), obținem rezultatul căutat.

Se poate vedea ușor că dacă $(\tilde{x}_n)$ converge în medie de ordinul r spre $\tilde{x}$ și $\tilde{y}$, atunci $\tilde{x} = \tilde{y}$ aproape sigur. Reciproc, dacă $(\tilde{x}_n)$ converge în medie de ordinul r spre $\tilde{x}$ și $\tilde{x} = \tilde{y}$ aproape sigur, atunci $(\tilde{x}_n)$ converge în medie de ordinul r spre $\tilde{y}$.

Se poate arăta de asemenea că spațiul semimetric $\{\mathfrak{L}^r(E, \mathcal{K}, P), \rho_r\}$ este complet.

Dintre spațiile $\mathfrak{L}^r(E, \mathcal{K}, P)$, cele mai importante sînt: $\mathfrak{L}^1(E, \mathcal{K}, P) = \mathfrak{L}$ și $\mathfrak{L}^2(E, \mathcal{K}, P) = \mathfrak{L}^2$. Convergența în $\mathfrak{L}^2$ este numită și *convergență în medie pătratică*. Dacă $(\tilde{x}_n) \subset \mathfrak{L}^2$ converge în medie pătratică spre $\tilde{x}$, vom scrie

$$\text{l.i.m}_n \, \tilde{x}_n = \tilde{x}.$$

7.4. CONVERGENȚA ÎN REPARTIȚIE

Să considerăm cîmpul borelian de evenimente $(R^m, \mathfrak{B}^m)$ și un șir de măsuri probabiliste* $(P_n)_{n \geqslant 0}$ pe acest cîmp.

Definiția 7.4.1. *Spunem că șirul* $(P_n)_{n \geqslant 1}$ *converge slab spre* P_0 *dacă pentru orice funcție continuă și mărginită* $g : R^m \to R$, *avem*

$$\lim_n \int_{R^m} g(x) \, P_n(dx) = \int_{R^m} g(x) \, P_0(dx) . \qquad (7.4.1)$$

Să considerăm un cîmp borelian de probabilitate $(E, \mathcal{K}, P)$ și un șir de variabile aleatoare $(\tilde{x}_n)_{n \geqslant 0}$ pe acest cîmp.

---

* Definiția care urmează poate fi dată pentru măsuri finite arbitrare.

Definiția 7.4.2. *Spunem că șirul de variabile aleatoare* $(\widetilde{x}_n)_{n\geq 1}$ *converge în repartiție spre* $\widetilde{x}_0$ *dacă* $P\circ\widetilde{x}_n^{-1}$ *converge slab spre* $P\circ\widetilde{x}_0^{-1}$.

Din această definiție rezultă că $(\widetilde{x}_n)$ converge în repartiție spre $\widetilde{x}_0$ dacă pentru orice funcție $g: R \to R$ mărginită și continuă avem

$$\lim_n \int_R g(x)\, dF_n(x) = \int_R g(x)\, dF_0(x), \tag{7.4.2}$$

unde F_n, $n \geq 0$, sînt funcțiile de repartiție ale variabilelor aleatoare $\widetilde{x}_n$, $n \geq 0$.

Se poate vedea ușor că dacă șirul de variabile aleatoare $(\widetilde{x}_n)$ converge în repartiție spre $\widetilde{x}$ și spre $\widetilde{y}$, atunci $\widetilde{x}$ și $\widetilde{y}$ au aceeași funcție de repartiție (sînt *identic repartizate*). Invers, dacă $(\widetilde{x}_n)$ converge în repartiție spre $\widetilde{x}$, iar $\widetilde{x}$ și $\widetilde{y}$ sînt identic repartizate, atunci $(\widetilde{x}_n)$ converge în repartiție spre $\widetilde{y}$.

Din cele de mai sus rezultă posibilitatea de a defini convergența în repartiție sau convergența slabă a șirurilor de funcții de repartiție.

Definiția 7.4.3. *Spunem că șirul de funcții de repartiție* $(F_n)_{n\geq 1}$ *converge în repartiție spre funcția de repartiție* F_0 *dacă pentru orice funcție* $g: R \to R$ *mărginită și continuă avem satisfăcută condiția* (7.4.2).

Avem rezultatul următor.

Teorema 7.4.1. *Fie* $(E, \mathcal{K}, P)$ *un cîmp borelian de probabilitate,* $(\widetilde{x}_n)_{n\geq 0}$ *un șir de variabile aleatoare pe acest cîmp și* $(F_n)_{n\geq 0}$ *șirul funcțiilor de repartiție corespunzătoare acestor variabile aleatoare. În aceste condiții, următoarele afirmații sînt echivalente:*

(a) Șirul $(\widetilde{x}_n)_{n\geq 1}$ *converge în repartiție spre* $\widetilde{x}_0$;

(b) Șirul $(F_n)_{n\geq 1}$ *converge în repartiție spre* F_0;

(c) Pentru orice punct de continuitate al funcției F_0,

$$\lim_n F_n(x) = F_0(x); \tag{7.4.3}$$

(d) Există o mulțime $A \subset R$ *densă astfel încît*

$$\lim_n F_n(x) = F_0(x) \tag{7.4.4}$$

pentru orice $x \in A$.

Demonstrație. Pentru a arăta echivalența afirmațiilor de mai sus vom demonstra lanțul de implicații $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow$ $\Rightarrow (d) \Rightarrow (a)$.

$(a) \Rightarrow (b)$. Este evidentă.

$(b) \Rightarrow (c)$. Fie x un punct de continuitate al lui F_0. Pentru orice $\varepsilon > 0$ există $\delta = \delta(x, \varepsilon) > 0$ astfel încît

$$| F_0(y) - F_0(x) | < \varepsilon \tag{7.4.5}$$

pentru orice y pentru care $| y - x | < \delta(x, \varepsilon)$.

Să considerăm funcțiile

$$g_1(y) = \begin{cases} 1 & \text{pentru } y \leqslant x - \delta, \\ (x - y)/\delta & \text{pentru } x - \delta \leqslant y \leqslant x, \\ 0 & \text{pentru } y \geqslant x \end{cases} \tag{7.4.6}$$

și

$$g_2(y) = \begin{cases} 1 & \text{pentru } y \leqslant x, \\ 1 - (y - x)/\delta & \text{pentru } x \leqslant y \leqslant x + \delta, \\ 0 & \text{pentru } y \geqslant x + \delta, \end{cases} \tag{7.4.7}$$

care sînt evident continue și mărginite pe R. Deci din ipoteză

$$\lim_n \int_R g_i(y)\, dF_n(y) = \int_R g_i(y)\, dF_0(y) \tag{7.4.8}$$

pentru $i = 1, 2$. Cu alte cuvinte, există un număr natural $N(\varepsilon)$ cu proprietatea că

$$\left| \int_R g_i(y)\, dF_n(y) - \int_R g_i(y)\, dF_0(y) \right| < \varepsilon, \quad i = 1, 2, \tag{7.4.9}$$

pentru $n \geqslant N(\varepsilon)$.

Pe de altă parte, avem evident

$$\begin{aligned} \int_R g_1(y)\, dF_0(y) &\geqslant \int_{-\infty}^{x-\delta} dF_0(y) = F_0(x - \delta), \\ \int_R g_2(y)\, dF_0(y) &\leqslant \int_{-\infty}^{x+\delta} dF_0(y) = F_0(x + \delta) \end{aligned} \tag{7.4.10}$$

și

$$\int_R g_1(y)\,dF_n(y) \leqslant \int_{-\infty}^{x} dF_n(y) = F_n(x),$$
$$\int_R g_2(y)\,dF_n(y) \geqslant \int_{-\infty}^{x} dF_n(y) = F_n(x). \tag{7.4.11}$$

Utilizînd relațiile (7.4.5), (7.4.9), (7.4.10) și (7.4.11) obținem

$$|F_n(x) - F_0(x)| \leqslant 2\varepsilon$$

pentru $n \geqslant N(\varepsilon)$, ceea ce demonstrează afirmația.

$(c) \Rightarrow (d)$. Implicația este evidentă, deoarece mulțimea punctelor de discontinuitate ale funcției F_0 este cel mult numărabilă și deci mulțimea punctelor de continuitate ale lui F_0 este densă în R.

$(d) \Rightarrow (a)$. Să considerăm o funcție continuă și mărginită $g : R \to R$ și fie

$$M = \sup_{x \in R} |g(x)|.$$

Pentru orice $\varepsilon > 0$ se pot determina numerele reale $a = a(\varepsilon)$ și $b = b(\varepsilon)$ din A, astfel încît

$$a < b,\ F_0(a) \leqslant \varepsilon,\ \ 1 - F_0(b) \leqslant \varepsilon. \tag{7.4.12}$$

Mai mult, din uniform continuitatea lui g pe $[a, b]$ rezultă că se pot determina punctele $a = x_0 < x_1 < \ldots < x_m = b$ cu $x_i \in A$, $0 \leqslant i \leqslant m$, astfel încît

$$|g(x) - g(x_i)| \leqslant \varepsilon \tag{7.4.13}$$

pentru orice $x \in [x_{i-1}, x_i]$, $1 \leqslant i \leqslant m$.

Dacă se consideră funcția

$$g_\varepsilon(x) = \begin{cases} g(x_{i-1}) & \text{pentru } x \in [x_{i-1}, x_i),\ 1 \leqslant i \leqslant m, \\ 0 & \text{în caz contrar,} \end{cases} \tag{7.4.14}$$

atunci, evident,

$$|g(x) - g_\varepsilon(x)| < \varepsilon \tag{7.4.15}$$

pentru orice $x \in [a, b]$.

Evident, pentru orice funcție de repartiție F, avem

$$\int_R g_\varepsilon(x)\, dF(x) = \sum_{i=1}^{m} g(x_{i-1})\,[F(x_i) - F(x_{i-1})]. \tag{7.4.16}$$

Deoarece $F_n(x_i) \to F_0(x_i)$, $0 \leqslant i \leqslant m$, din (7.4.16) rezultă că

$$\lim_n \int_R g_\varepsilon(x)\, dF_n(x) = \int_R g_\varepsilon(x)\, dF_0(x).$$

Există deci numărul natural $N_1(\varepsilon)$ astfel încît

$$\left| \int_R g_\varepsilon(x)\, dF_n(x) - \int_R g_\varepsilon(x)\, dF_0(x) \right| \leqslant \varepsilon \tag{7.4.17}$$

pentru orice $n \geqslant N_1(\varepsilon)$.

Pe de altă parte, pentru orice funcție de repartiție F, are loc

$$\begin{aligned} \int_R |g(x) - g_\varepsilon(x)|\, dF(x) &= \int_{-\infty}^{a} |g(x) - g_\varepsilon(x)|\, dF(x) + \\ + \int_a^b |g(x) - g_\varepsilon(x)|\, dF(x) &+ \int_b^{\infty} |g(x) - g_\varepsilon(x)|\, dF(x) \leqslant \\ &\leqslant MF(a) + \varepsilon(F(b) - F(a)) + M(1 - F(b)). \end{aligned} \tag{7.4.18}$$

Înlocuind F cu F_0 în (7.4.18) și ținînd seama de (7.4.12), obținem

$$\int_R |g(x) - g_\varepsilon(x)|\, dF_0(x) \leqslant (2M + 1)\,\varepsilon. \tag{7.4.19}$$

Scanned with OKEN Scanner

Deoarece $a, b \in A$, rezultă că $F_n(a) \to F_0(a)$ și $F_n(b) \to F_0(b)$ și deci există $N_2(\varepsilon)$ astfel încît

$$| F_n(a) - F_0(a) | \leqslant \varepsilon, \quad | F_n(b) - F_0(b) | \leqslant \varepsilon \tag{7.4.20}$$

pentru $n \geqslant N_2(\varepsilon)$.

Înlocuind în (7.4.18) pe F prin F_n și ținînd seama de (7.4.12) și (7.4.20), obținem

$$\int_R | g(x) - g_\varepsilon(x) | \, dF_n(x) \leqslant (4M + 1)\, \varepsilon \tag{7.4.21}$$

pentru orice $n \geqslant N_2(\varepsilon)$.

Fie $N(\varepsilon) = \max(N_1(\varepsilon), N_2(\varepsilon))$. Din (7.4.17), (7.4.19) și (7.4.21) rezultă că pentru $n \geqslant N(\varepsilon)$ are loc

$$\left| \int_R g(x)\, dF_n(x) - \int_R g_\varepsilon(x)\, dF_0(x) \right| \leqslant \int_R | g(x) - g_\varepsilon(x) | \, dF_n(x) +$$

$$+ \int_R | g(x) - | g_\varepsilon(x) | \, dF_0(x) + \left| \int_R g_\varepsilon(x)\, dF_n(x) - \int_R g_\varepsilon(x)\, dF_0(x) \right| \leqslant$$

$$\leqslant 3(2M + 1)\, \varepsilon,$$

ceea ce demonstrează afirmația.

Legătura dintre convergența în repartiție și celelalte tipuri de convergență introduse mai sus este precizată de următoarele două teoreme.

Teorema 7.4.2. *Dacă șirul de variabile aleatoare* $(\tilde{x}_n)$ *converge în probabilitate spre* $\tilde{x}_0$, *atunci șirul considerat converge în repartiție spre* $\tilde{x}_0$.

Demonstrație. Fie $(F_n)_{n \geqslant 0}$ șirul funcțiilor de repartiție corespunzătoare variabilelor aleatoare $(\tilde{x}_n)_{n \geqslant 0}$ și fie $x \in R$ un punct de continuitate al lui F_0. Conform teoremei precedente rezultă că va fi suficient să arătăm că $F_n(x) \rightarrow F_0(x)$.

Deoarece x este punct de continuitate al lui F_0, pentru orice $\varepsilon > 0$ se poate determina $\delta = \delta(x, \varepsilon)$ astfel încît

$$| F_0(y) - F_0(x) | < \varepsilon \tag{7.4.22}$$

pentru orice y pentru care $|y - x| < \delta$. Să observăm însă că

$$\{\omega : \tilde{x}_0(\omega) < x - \delta/2\} = \{\omega : \tilde{x}_0(\omega) < x - \delta/2\} \cap$$
$$\cap (\{\omega : \tilde{x}_n(\omega) < x\} \cup \{\omega : \tilde{x}_n(\omega) \geqslant x\}) \subset$$
$$\subset \{\omega : \tilde{x}_n(\omega) < x\} \cup \{\omega : \tilde{x}_n(\omega) - \tilde{x}_0(\omega) > \delta/2\} \subset$$
$$\subset \{\omega : \tilde{x}_n(\omega) < x\} \cup \{\omega : |\tilde{x}_n(\omega) - \tilde{x}_0(\omega)| \geqslant \delta/2\}.$$

și deci

$$P(\tilde{x}_0 < x - \delta/2) \leqslant P(\tilde{x}_n < x) + P(|\tilde{x}_n - \tilde{x}_0| \geqslant \delta/2)$$

sau

$$F_0(x - \delta/2) \leqslant F_n(x) + P(|\tilde{x}_n - \tilde{x}_0| \geqslant \delta/2). \quad (7.4.23)$$

Ținînd seama că $(\tilde{x}_n)$ converge în probabilitate spre $\tilde{x}_0$, din (7.4.23) obținem prin trecere la limită

$$F_0(x - \delta/2) \leqslant \underline{\lim}_n F_n(x). \quad (7.4.24)$$

Schimbînd în (7.4.23) locul funcțiilor F_0 și F_n și înlocuind x prin $x + \delta/2$ obținem

$$F_n(x) \leqslant F_0(x + \delta/2) + P(|\tilde{x}_n - \tilde{x}_0| \geqslant \delta/2) \quad (7.4.25)$$

și deci

$$\overline{\lim}_n F_n(x) \leqslant F_0(x + \delta/2). \quad (7.4.26)$$

Din (7.4.24), (7.4.26) și ținînd seama de (7.4.22) obținem

$$0 \leqslant \overline{\lim}_n F_n(x) - \underline{\lim}_n F_n(x) \leqslant F_0(x + \delta/2) - F_0(x - \delta/2) < \varepsilon,$$

ceea ce arată că $(F_n(x))$ este un șir convergent.

Pentru a termina demonstrația mai trebuie arătat că

$$\lim_n F_n(x) = F_0(x). \quad (7.4.27)$$

Într-adevăr, din (7.4.22) și (7.4.24) rezultă

$$\lim_n F_n(x) - F_0(x) \geqslant F_0(x - \delta/2) - F_0(x) > \varepsilon,$$

Scanned with OKEN Scanner

iar din (7.4.22) și (7.4.26) obținem

$$\lim_n F_n(x) - F_0(x) \leqslant F_0(x + \delta/2) - F_0(x) < \varepsilon.$$

Ultimele două relații implică (7.4.27) și teorema este astfel demonstrată.

Observație. Conform observației de la teorema 7.2.3, pentru orice funcție continuă $g : R \to R$, rezultă că $(g(\tilde{x}_n))$ este un șir de variabile aleatoare convergent în probabilitate spre $g(\tilde{x}_0)$. Dacă în plus g este mărginită, atunci $g(\tilde{x}_n)$ sînt variabile aleatoare mărginite uniform și, din teorema convergenței dominate a lui Lebesgue*, rezultă că

$$\lim_n \int_E g(\tilde{x}_n)\, dP = \int_E g(\tilde{x}_0)\, dP,$$

ceea ce este echivalent cu convergența slabă a lui $P \circ \tilde{x}_n^{-1}$ spre $P \circ \tilde{x}_0^{-1}$. Așadar, avem o nouă demonstrație pentru teorema 7.4.2.

Reciproca acestei teoreme nu este adevărată. Într-adevăr, să considerăm cîmpul borelian de probabilitate $(E, \mathfrak{K}, P)$, în care $E = [0, 1]$, $\mathfrak{K}$ este corpul borelian corespunzător, iar P este măsura Lebesgue. Să considerăm șirul de variabile aleatoare $(\tilde{x}_n)_{n \geqslant 0}$ definite de

$$\tilde{x}_0(\omega) = \begin{cases} 1 & \text{pentru } \omega \in [0,\ 1/2], \\ 0 & \text{pentru } \omega \in (1/2,\ 1] \end{cases}$$

și

$$\tilde{x}_n(\omega) = \begin{cases} 0 & \text{pentru } \omega \in [0,\ 1/2], \\ 1 & \text{pentru } \omega \in (1/2,\ 1] \end{cases}$$

pentru $n \geqslant 1$.

Este clar că $(\tilde{x}_n)_{n \geqslant 0}$ sînt identic repartizate și deci $(\tilde{x}_n)$ converge în repartiție spre $\tilde{x}_0$. Pe de altă parte, avem

$$P(\{\omega : |\, \tilde{x}_n(\omega) - \tilde{x}_0(\omega)\, | = 1) = P([0,\ 1]) = 1,$$

ceea ce arată că $(\tilde{x}_n)$ nu converge în probabilitate spre $\tilde{x}_0$.

---

* Vezi HALMOS, P. R. (1950).

Teorema care urmează precizează condițiile în care afirmația reciprocă are loc.

Teorema 7.4.3. *Dacă șirul de variabile aleatoare* $(\widetilde{x}_n)$ *pe cîmpul de probabilitate* $(E, \mathfrak{K}, P)$ *converge în repartiție spre* $\widetilde{x}_0$ *și dacă* $\widetilde{x}_0$ *este constantă aproape sigur, atunci* $(\widetilde{x}_n)$ *converge în probabilitate spre* $\widetilde{x}_0$.

Demonstrație. Dacă $\widetilde{x}_0$ este constantă aproape sigur, atunci există un număr real x_0 astfel încît

$$P(\{\omega : \widetilde{x}_0(\omega) = x_0\}) = 1.$$

Funcția de repartiție F_0 a lui $\widetilde{x}_0$ este deci

$$F_0(x) = \begin{cases} 0 & \text{pentru } x \leqslant x_0, \\ 1 & \text{pentru } x > x_0, \end{cases}$$

ceea ce arată că x_0 este singurul punct de discontinuitate al lui F_0.

Dacă notăm prin F_n funcția de repartiție a lui $\widetilde{x}_n$, atunci pentru orice $\varepsilon > 0$ avem

$$P(|\widetilde{x}_n - x_0| \geqslant \varepsilon) \leqslant P(\widetilde{x}_n \leqslant x_0 - \varepsilon) + P(\widetilde{x}_n \geqslant x_0 + \varepsilon) =$$

$$= F_n(x_0 - \varepsilon + 0) + 1 - F_n(x_0 + \varepsilon) \leqslant F_n(x_0 - \varepsilon/2) + 1 - F_n(x_0 + \varepsilon).$$

Deoarece $x_0 - \varepsilon/2$ și $x_0 + \varepsilon/2$ sînt puncte de continuitate ale lui F_0, trecînd la limită în inegalitatea precedentă, obținem

$$\overline{\lim_n} P(|\widetilde{x}_n - x_0| \geqslant \varepsilon) \leqslant F_0(x_0 - \varepsilon/2) + 1 - F_0(x_0 + \varepsilon) =$$

$$= 0 + 1 - 1 = 0.$$

Deoarece avem, evident,

$$\underline{\lim_n} P(|\widetilde{x}_n - x_0| \geqslant \varepsilon) = 0,$$

rezultă

$$\lim_n P(|\widetilde{x}_n - x_0| \geqslant \varepsilon) = 0$$

și deci $(\widetilde{x}_n)$ converge în probabilitate spre $\widetilde{x}_0$.

Scanned with OKEN Scanner

7.5. COMPARAREA TIPURILOR DE CONVERGENȚĂ

În mulțimea variabilelor aleatoare definite pe un cîmp borelian de probabilitate am introdus diverse tipuri de convergență a șirurilor de variabile aleatoare: convergența aproape sigură, convergența tare, convergența în probabilitate, convergența în medie de ordinul r și convergența în repartiție. Legăturile dintre aceste tipuri de convergență au fost precizate în §§ 5.1—5.4 și pot fi rezumate în modul arătat în fig. 7.7.1. Din această figură se observă că convergența aproape sigură și convergența în medie de ordinul r sînt cele mai tari moduri de convergență, iar convergența în repartiție este cel mai slab mod de convergență.

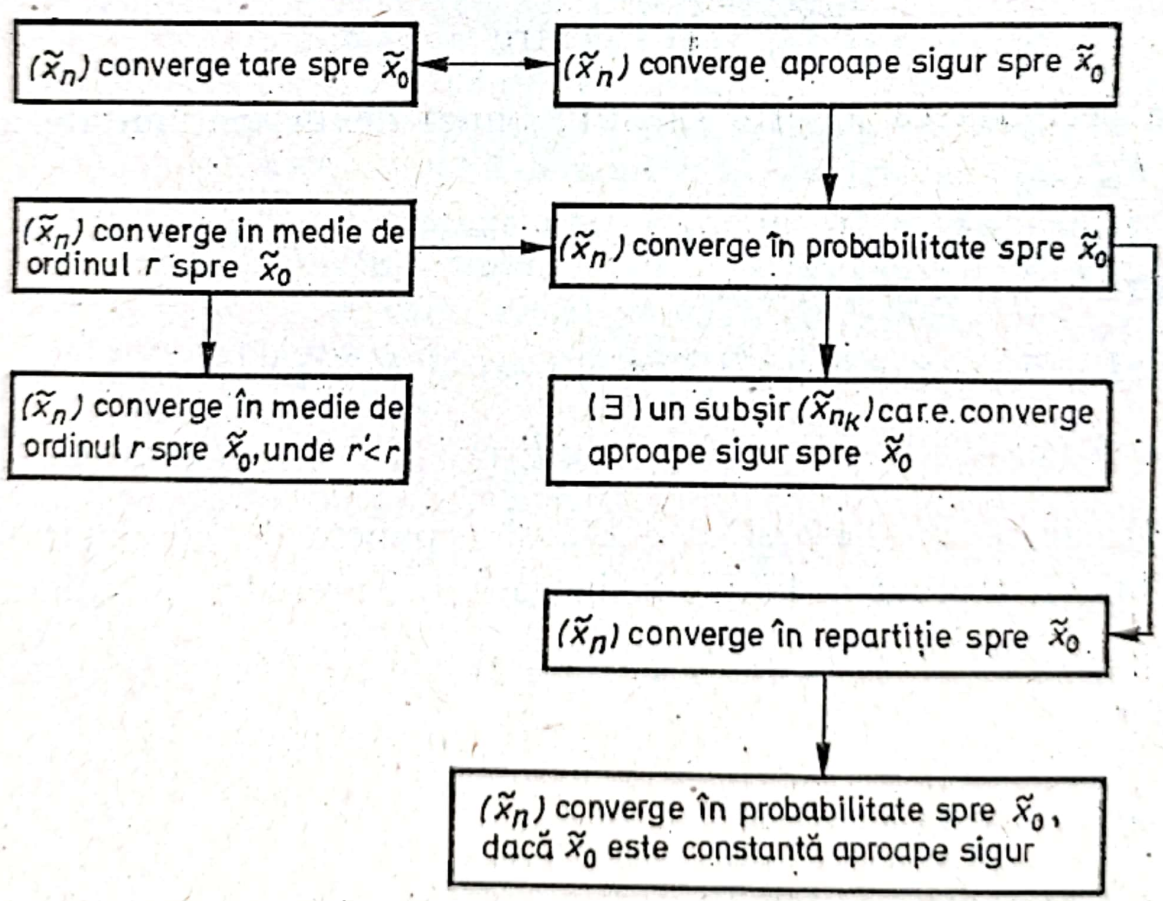


Fig. 7.7.1

Scanned with OKEN Scanner

Capitolul 8

FUNCȚII CARACTERISTICE

8.1. INTRODUCERE

În capitolele precedente am asociat unei variabile aleatoare sau unui vector aleator funcțiile de repartiție corespunzătoare. Instrumentul analitic furnizat de funcția de repartiție este în general destul de greu de manipulat. De exemplu, sumei unor variabile aleatoare independente îi corespunde produsul de compoziție al funcțiilor de repartiție ale termenilor, care este dificil de calculat în multe cazuri.

Această situație a impus introducerea unei alte funcții care să caracterizeze variabila sau vectorul aleator și care să fie mai ușor de manipulat; acest nou instrument analitic este furnizat de *funcția caracteristică*. Datorită faptului că funcția caracteristică a unei sume de variabile aleatoare independente este produsul funcțiilor caracteristice ale termenilor, funcția caracteristică este în multe cazuri un instrument analitic mult mai util decît funcția de repartiție.

Să considerăm un cîmp borelian de probabilitate $(E, \mathfrak{K}, P)$ și fie $\tilde{x}$ și $\tilde{y}$ două variabile aleatoare pe acest cîmp. Funcția cu valori complexe

$$\tilde{z}(\omega) = \tilde{x}(\omega) + \mathrm{i}\tilde{y}(\omega) \tag{8.1.1}$$

este numită *variabila aleatoare complexă*. Dacă $\tilde{x}$ și $\tilde{y}$ sînt integrabile, atunci $\tilde{z}$ este de asemenea numită integrabilă și, în plus, avem

$$\int_E \tilde{z}(\omega)\,\mathrm{d}P(\omega) = \int_E \tilde{x}(\omega)\,\mathrm{d}P(\omega) + \mathrm{i}\int_E \tilde{y}(\omega)\,\mathrm{d}P(\omega). \tag{8.1.2}$$

Rezultă deci că dacă $\tilde{z}$ este integrabilă, atunci media sa există și este dată de relația

$$M(\tilde{z}) = M(\tilde{x}) + \mathrm{i}\,M(\tilde{y}). \tag{8.1.3}$$

Scanned with OKEN Scanner

Pentru orice variabilă aleatoare $\tilde{x}$ cu valori reale, funcţia

$$e^{i t\tilde{x}(\omega)} = \cos(t\tilde{x}(\omega)) + i\sin(t\tilde{x}(\omega)) \tag{8.1.4}$$

este, pentru fiecare t real, o variabilă aleatoare complexă mărginită (modulul său este egal cu 1) și deci integrabilă.

Definiția 8.1.1. *Funcția de variabilă reală* $\varphi_{\tilde{x}}(t)$ *definită de*

$$\varphi_{\tilde{x}}(t) = M(e^{it\tilde{x}}) \tag{8.1.5}$$

este numită funcția caracteristică a variabilei aleatoare $\tilde{x}$.

Cînd nu există pericol de confuzie indicele $\tilde{x}$ se suprimă.

Evident, dacă F este funcția de repartiție a variabilei aleatoare $\tilde{x}$, atunci

$$\varphi_{\tilde{x}}(t) = \int_R e^{itx}\, dF(x), \tag{8.1.6}$$

ceea ce arată că funcția caracteristică este complet determinată de funcția de repartiție. Așa cum vom vedea în cele ce urmează, este adevărată și afirmația reciprocă: funcția de repartiție este complet determinată de funcția caracteristică.

Este clar că dacă $\tilde{x}$ are densitatea de probabilitate $f(x)$, atunci funcția sa caracteristică este

$$\varphi_{\tilde{x}}(t) = \int_R e^{itx} f(x)\, dx. \tag{8.1.7}$$

Analog, dacă $\tilde{x}$ este discretă și ia valorile $(x_j)_{j\in I}$, atunci funcția sa caracteristică este

$$\varphi_{\tilde{x}}(t) = \sum_{j\in I} e^{itx_j} f(x_j), \tag{8.1.8}$$

unde f este funcția de masă a variabilei aleatoare $\tilde{x}$.

Să considerăm vectorul aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)'$ pe $(E, \mathcal{K}, P)$.

Definiția 8.1.2. *Funcția de* n *variabile reale* $\varphi_{\tilde{\mathbf{x}}}(t_1, \ldots, t_n)$ *definită de*

$$\varphi_{\tilde{\mathbf{x}}}(t_1, \ldots, t_n) = M\left(e^{i\sum_{j=1}^{n} t\tilde{(x)}}\right) \tag{8.1.9}$$

este numită funcția caracteristică a vectorului aleator $\tilde{\mathbf{x}}$.

Evident, dacă $F(x_1, ..., x_n)$ este funcția de repartiție a vectorului aleator $\tilde{x}$, atunci

$$\varphi_{\tilde{x}}(t_1, ..., t_n) = \int_{R^n} e^{i(t_1x_1 + ... + t_nx_n)} \, dF(x_1, ..., x_n). \qquad (8.1.10)$$

8.2. PROPRIETĂȚILE FUNCȚIEI CARACTERISTICE

Să considerăm un cîmp borelian de probabilitate $(E, \mathcal{K}, P)$ și un vector aleator $\tilde{x} = (\tilde{x}_1, ..., \tilde{x}_n)'$ pe acest cîmp, a cărui funcție de repartiție este $F(x_1, ..., x_n)$.

Teorema 8.2.1. *Funcția caracteristică* $\varphi_{\tilde{x}}(t_1, ..., t_n)$ *a vectorului aleator* $\tilde{x} = (\tilde{x}_1, ..., \tilde{x}_n)'$ *are următoarele proprietăți:*

(*a*) $\varphi_{\tilde{x}}(0, ..., 0) = 1$;

(*b*) $|\varphi_{\tilde{x}}(t_1, ... t_n)| \leqslant 1$;

(*c*) $\varphi_{\tilde{x}}(-t_1, ..., -t_n) = \overline{\varphi_{\tilde{x}}(t_1, ..., t_n)}$;

(*d*) $\varphi_{\tilde{x}}(t_1, ..., t_n)$ *este uniform continuă pe* R^n;

(*e*) *Dacă există momentul de ordinul* $(r_1, ..., r_n)'$ *al vectorului aleator* $\tilde{x}$ *și* $r = r_1 + ... + r_n$, *atunci*

$$M_{r_1 ... r_n}(\tilde{x}) = i^r \left[\frac{\partial^r \varphi_{\tilde{x}}(t_1, ..., t_n)}{\partial^{r_1} t_1 ... \partial^{r_n} t_n} \right]_{t_1 = ... = t_n = 0}; \qquad (8.2.1)$$

(*f*) *Dacă* $\tilde{y}(\omega) = (a_1\tilde{x}_1(\omega) + b_1, ..., a_n\tilde{x}_n(\omega) + b_n)'$ *cu* $a_j, b_j \in R$, $1 \leqslant j \leqslant n$, *atunci*

$$\varphi_{\tilde{y}}(t_1, ..., t_n) = \varphi_{\tilde{x}}(a_1t_1, ..., a_nt_n) \, e^{i \sum_{j=1}^{n} b_jt_j}; \qquad (8.2.2)$$

(*g*) *Dacă* $\tilde{y}(\omega) = \sum_{j=1}^{n} a_j\tilde{x}_j(\omega)$ *cu* $a_j \in R$, $1 \leqslant j \leqslant n$, *atunci*

$$\varphi_{\tilde{y}}(t) = \varphi_{\tilde{x}}(ta_1, ..., ta_n); \qquad (8.2.3)$$

(*h*) *Dacă* $\tilde{x}_1, ..., \tilde{x}_n$ *sînt variabile aleatoare independente și*

$$\tilde{y}(\omega) = \tilde{x}_1(\omega) + ... + \tilde{x}_n(\omega),$$

atunci

$$\varphi_{\tilde{y}}(t) = \varphi_{\tilde{x}_1}(t) ... \varphi_{\tilde{x}_n}(t); \qquad (8.2.4)$$

Scanned with OKEN Scanner

(*i*) *Pentru orice* m *natural, orice numere complexe* $z_1, \ldots, z_m$ *și orice puncte* $t^1, \ldots, t^m \in R^n$, *avem*

$$\sum_{j,k=1}^{m} \varphi_{\tilde{x}}(t_1^j - t_1^k, \ldots, t_n^j - t_n^k) z_j \overline{z_k} \geqslant 0. \tag{8.2.5}$$

Demonstrație. (*a*), (*b*) și (*c*) sînt evidente.

(*d*) Pentru simplitatea scrierii, vom da demonstrația în cazul $n = 1$. Avem

$$|\varphi(t+h) - \varphi(t)| = \left| \int_R (e^{i(t+h)} - e^{itx})\, dF(x) \right| \leqslant$$

$$\leqslant \int_R |e^{itx}| \cdot |e^{ihx} - 1|\, dF(x) = \int_R |e^{i \cdot hx} - 1|\, dF(x).$$

Să observăm însă că $e^{ihx} \to 1$ pentru $h \to 0$ și x fixat, iar $|e^{ihx} - 1| \leqslant 2$ și deci

$$\lim_{h \to 0} \int_R |e^{ihx} - 1|\, dF(x) = 0$$

din teorema convergenței dominate a lui Lebesgue. Rezultă că pentru orice $\varepsilon > 0$ există $\delta(\varepsilon)$ astfel încît

$$|\varphi(t+h) - \varphi(t)| < \varepsilon$$

pentru orice h pentru care $|h| < \delta(\varepsilon)$, de unde rezultă uniform continuitatea lui φ.

(*e*) Derivînd de r ori în (8.1.10) obținem

$$\frac{\partial^r \varphi(t_1, \ldots, t_n)}{\partial t_1^{r_1} \ldots \partial t_n^{r_n}} = \int_{R^n} i^r t_1^{r_1} \ldots t_n^{r_n} e^{i(t_1 x_1 + \ldots + t_n x_n)}\, dF(x_1, \ldots, x_n). \tag{8.2.6}$$

Avem însă

$$\left| \int_{R^n} t_1^{r_1} \ldots t_n^{r_n}\, dF(x_1, \ldots, x_n) \right| \leqslant \int_{R^n} |t_1^{r_1} \ldots t_n^{r_n}|\, dF(x_1, \ldots, x_n) = M(\tilde{x}_1^{r_1} \ldots \tilde{x}_n^{r_n})$$

și, conform ipotezei, această expresie este mărginită, ceea ce arată că integrala din (8.2.6) există. Punînd în (6.2.6) $t_1 = \ldots = t_n = 0$ se obține formula căutată (8.2.1).

Scanned with OKEN Scanner

(*f*) Avem

$$\varphi_{\tilde{y}}(t_1, \ldots, t_n) = M\left(e^{i\sum_{j=1}^{n} t_j(a_j\tilde{x}_j(\omega)+b_j)}\right) = M\left(e^{i\sum_{j=1}^{n} t_j b_j} \cdot e^{i\sum_{j=1}^{n} a_j t_j \tilde{x}_j(\omega)}\right) =$$

$$= \varphi_{\tilde{x}}(a_1 t_1, \ldots, a_n t_n)\, e^{i\sum_{j=1}^{n} b_j t_j}.$$

(*g*) Avem

$$\tilde{y}(t) = M\left(e^{it\sum_{j=1}^{n} a_j\tilde{x}_j(\omega)}\right) = M\left(e^{i\sum_{j=1}^{n} t a_j x_j(\omega)}\right) = \varphi_{\tilde{x}}(a_1 t, \ldots, a_n t).$$

(*h*) Avem

$$\varphi_{\tilde{y}}(t) = M\left(e^{it\sum_{j=1}^{n}\tilde{x}_j(\omega)}\right) = M(e^{it\tilde{x}_1(\omega)} \ldots e^{it\tilde{x}_n(\omega)}) =$$

$$= M(e^{it\tilde{x}_1}) \ldots M(e^{it\tilde{x}_n}) = \varphi_{\tilde{x}_1}(t) \ldots \varphi_{\tilde{x}_n}(t),$$

deoarece $e^{it\tilde{x}_1}, \ldots, e^{it\tilde{x}_n}$ sînt variabile aleatoare independente.

(*i*) Avem

$$\sum_{j,k=1}^{m} \varphi_{\tilde{x}}(t_1^j - t_1^k, \ldots, t_n^j - t_n^k)\, z_j \overline{z}_k =$$

$$= \int_{R^n} \left(\sum_{j,k=1}^{m} e^{i\sum_{r=1}^{n} (t_r^j - t_r^k) x_r} z_j \overline{z}_k \right) dF(x_1, \ldots, x_n) =$$

$$= \int_{R_n} \left(\sum_{j=1}^{m} e^{i\sum_{r=1}^{n} t_r^j x_r} z_j \right) \left(\sum_{k=1}^{n} e^{-i\sum_{r=1}^{n} t_r^k x_r} \overline{z}_k \right) dF(x_1, \ldots, x_n) =$$

$$= \int_{R^n} \left| \sum_{j=1}^{m} e^{i\sum_{r=1}^{n} t_r^j x_r} z_j \right|^2 dF(x_1, \ldots, x_n) \geqslant 0,$$

ceea ce demonstrează afirmația.

Observația 1. Proprietatea $(h)''$ este fundamentală în aplicații, permițînd calculul simplu al funcției caracteristice a sumei de variabile aleatoare independente.

Observația 2. O funcție care verifică condiția (8.2.5) este numită și *funcție de tip pozitiv*. Conform teoremei precedente rezultă că funcția caracteristică este o funcție de tip pozitiv. Se poate demonstra și o teoremă reciprocă (teorema lui Bochner), care precizează condițiile în care o funcție de tip pozitiv este funcția caracteristică a unei anumite variabile aleatoare.

8.3. FORMULA DE INVERSIUNE

Am văzut că funcția de repartiție a unei variabile aleatoare determină funcția caracteristică. Problema care se pune este dacă are loc și rezultatul reciproc. Răspunsul la această problemă este furnizat de teorema următoare.

Teorema 8.3.1. *Fie* $\tilde{x}$ *o variabilă aleatoare pe cîmpul borelian de probabilitate* $(E, \mathcal{K}, P)$ *și fie* F *și* φ *respectiv funcția de repartiție și funcția caracteristică corespunzătoare acestei variabile aleatoare. În aceste condiții, pentru orice puncte de continuitate* $x_1, x_2 \in R$ *ale funcției* F *cu* $x_1 < x_2$, *avem*

$$F(x_2) - F(x_1) = \lim_{c \to +\infty} \frac{1}{2\pi} \int_{-c}^{c} \frac{e^{-itx_1} - e^{-itx_2}}{it} \varphi(t)\, dt, \qquad (8.3.1)$$

formulă cunoscută sub numele de formula de inversiune.

Demonstrație. Să observăm mai întîi că integrala

$$I(c) = \frac{1}{2\pi} \int_{-c}^{c} \frac{e^{-itx_1} - e^{-itx_2}}{it} \varphi(t)\, dt \qquad (8.3.2)$$

este o integrală riemanniană obișnuită, funcția de sub semnul integrală fiind continuă și mărginită pe R (în $t = 0$ valoarea funcției este definită prin continuitate ca fiind egală cu $(x_2 - x_1)\,\varphi(0)$, iar funcția este mărginită pe R chiar de această

valoare). Ținînd seama de definiția funcției caracteristice, rezultă

$$I(c) = \frac{1}{2\pi}\int_{-c}^{c}\int_{R}\frac{e^{it(x-x_1)} - e^{it(x-x_2)}}{it}\,dF(x)\,dt. \qquad (8.3.3)$$

Deoarece integrala în raport cu x este absolut convergentă, iar integrala în raport cu t se face pe interval finit, în (8.3.3) ordinea de integrare poate fi inversată și obținem

$$I(c) = \frac{1}{2\pi}\int_{R}\left[\int_{-c}^{c}\frac{e^{it(x-x_1)} - e^{it(x-x_2)}}{it}\,dt\right]dF(x) =$$

$$= \frac{1}{2\pi}\int_{R}\left[\int_{-c}^{c}\left(\frac{e^{it(x-x_1)} - e^{-it(x-x_1)}}{it} - \frac{e^{it(x-x_2)} - e^{-it(x-x_2)}}{it}\right)dt\right]dF(x) =$$

$$= \frac{1}{\pi}\int_{R}\left[\int_{0}^{c}\frac{1}{t}\,(\sin t(x-x_1) - \sin t(x-x_2))\,dt\right]dF(x).$$

Dacă notăm

$$J(c, x) = \frac{1}{\pi}\int_{0}^{c}\frac{1}{t}\,(\sin t(x-x_1) - \sin t(x-x_2))\,dt =$$

$$= \frac{1}{\pi}\int_{c(x-x_2)}^{c(x-x_1)}\frac{\sin y}{y}\,dy \qquad (8.3.4)$$

rezultă deci că putem scrie

$$I(c) = \int_{R} J(c, x)\,dF(x). \qquad (8.3.5)$$

Ținînd seama de formula lui Dirichlet

$$\lim_{\substack{a\to-\infty\\ b\to+\infty}} \int_a^b \frac{\sin x}{x} = \pi,$$

rezultă

$$\lim_{c\to+\infty} J(c, x) = \begin{cases} 0 & \text{pentru } x\in(-\infty, x_1)\cup(x_2,+\infty), \\ 1/2 & \text{pentru } x = x_1,\ x_2, \\ 1 & \text{pentru } x\in(x_1, x_2). \end{cases} \tag{8.3.6}$$

Deoarece $J(c, x)$ este uniform mărginită ca funcție de c, din (8.3.5) și (8.3.6) rezultă

$$\lim_{c\to+\infty} I(c) = \int_R [\lim_{c\to+\infty} J(c, x)]\, dF(x) = \frac{1}{2}[F(x_1+0) - F(x_1-0)] +$$

$$+[F(x_2-0) - F(x_1+0)] + \frac{1}{2}[F(x_2+0) - F(x_2-0)]$$

sau

$$\lim_{c\to+\infty} I(c) = \frac{F(x_2-0)+F(x_2+0)}{2} \quad \frac{F(x_1-0)+F(x_1+0)}{2}. \tag{8.3.7}$$

Deoarece x_1 și x_2 sînt puncte de continuitate pentru F, formula (8.3.2) se deduce imediat din (8.3.7).

Observație. Din teorema precedentă rezultă că în fiecare punct de continuitate al funcției de repartiție F avem

$$F(x) = \lim_{y\to-\infty} \lim_{c\to\infty} \frac{1}{2\pi} \int_{-c}^{c} \frac{e^{-ity} - e^{-itx}}{it} \varphi(t)\, dt, \tag{8.3.8}$$

unde limita în raport cu y se ia cu condiția suplimentară ca y să fie punct de continuitate al lui F. Formula (8.3.8) arată că funcția de repartiție F este unic determinată de funcția caracteristică φ.

Scanned with OKEN Scanner

Dacă în plus F este derivabilă în punctul x, atunci derivata sa $f(x)$ este dată de relația

$$f(x) = \lim_{h\to 0} \lim_{c\to\infty} \frac{1}{2\pi} \int_{-c}^{c} \frac{1 - e^{-ith}}{ith} e^{-itx} \varphi(t)\, dt, \qquad (8.3.9)$$

cu condiția ca membrul drept din această formulă să existe.

În cazul cînd φ este absolut integrabilă pe R, formula (8.3.9) poate fi precizată astfel.

Teorema 8.3.2. *Dacă funcția caracteristică* φ *este absolut integrabilă pe* R, *atunci funcția de repartiție corespunzătoare* F *este absolut continuă, iar derivata sa* f(x) *este dată de relația*

$$f(x) = \frac{1}{2\pi} \int_{R} e^{-itx} \varphi(t)\, dt. \qquad (8.3.10)$$

Demonstrație. Dacă φ este absolut integrabilă, formula de inversiune (8.3.1) poate fi scrisă sub forma

$$F(x_2) - F(x_1) = \frac{1}{2\pi} \int_{R} \frac{e^{-itx_1} - e^{-itx_2}}{it} \varphi(t)\, dt,$$

unde $x_1 < x_2$ sînt puncte de continuitate ale lui F. Dacă punem $x_1 = x - h$, $x_2 = x + h$, rezultă că

$$F(x+h) - F(x-h) = \frac{2h}{2\pi} \int_{R} \frac{\sin th}{th} e^{-itx} \varphi(t)\, dt \leqslant \frac{2h}{2\pi} \int_{R} |\varphi(t)|\, dt,$$

de unde rezultă absolut continuitatea lui F. Din această relație obținem prin trecere la limită:

$$f(x) = \lim_{h\to 0} \frac{F(x+h) - F(x-h)}{2h} = \lim_{h\to 0} \frac{1}{2\pi} \int_{R} \frac{\sin th}{th} e^{-itx} \varphi(t)\, dt =$$

$$= \frac{1}{2\pi} \int_{R} e^{-itx} \varphi(t)\, dt,$$

ultima limită existînd conform teoremei lui Lebesgue de trecere la limită sub semnul integralei.

Scanned with OKEN Scanner

În sfîrșit, pentru a arăta că $f(x)$ este continuă, să observăm mai întîi că

$$|f(x+h)-f(x)| \leqslant (1/\pi)\int_R |\sin(th/2)| \cdot |\varphi(t)|\,dt =$$

$$= (1/\pi)\int_{|t|\leqslant a} |\sin(th/2)| \cdot |\varphi(t)|\,dt + (1/\pi)\int_{|t|>a} |\sin(th/2)| \cdot |\varphi(t)|\,dt,$$

unde $a > 0$ este un număr real arbitrar.

Dacă $\varepsilon > 0$ este dat, prima integrală poate fi majorată cu $\varepsilon/2$ dacă se alege $|h|$ suficient de mic, iar a este fixat. Analog, pentru h fixat, putem alege a suficient de mare pentru ca a doua integrală să fie mai mică decît $\varepsilon/2$. În aceste condiții avem

$$|f(x+h)-f(x)| < \varepsilon$$

pentru $|h|$ suficient de mic, ceea ce încheie demonstrația.

Rezultatele obținute mai sus se generalizează ușor la vectori aleatori. De exemplu, formula de inversiune capătă forma precizată în teorema următoare.

Teorema 8.3.3. *Fie* $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)'$ *un vector aleator pe cîmpul borelian de probabilitate* $(E, \mathcal{K}, P)$ *și fie* $F(x_1, \ldots, x_n)$ *și* $\varphi(t_1, \ldots, t_n)$ *respectiv funcția de repartiție și funcția caracteristică corespunzătoare acestui vector aleator. Dacă* $a_j < b_j$, $1 \leqslant j \leqslant n$, *sînt numere reale arbitrare cu proprietatea că probabilitatea ca* $\tilde{\mathbf{x}}$ *să se afle pe suprafața paralelipipedului*

$$\{\mathbf{x} : \mathbf{x} = (x_1, \ldots, x_n)', \quad a_j \leqslant x_j < b_j, \quad 1 \leqslant j \leqslant n\}$$

să fie nulă, atunci are loc următoarea formulă de inversiune

$$P(a_1 \leqslant \tilde{x}_1 < b_1, \ldots, a_n \leqslant \tilde{x}_n < b_n) =$$

$$= (2\pi)^{-n} \lim_{c \to -\infty} \int_{-c}^{c} \cdots \int_{-c}^{c} \prod_{j=1}^{n} \frac{e^{-it_j a_j} - e^{-it_j b_j}}{it_j} \varphi(t_1, \ldots, t_n)\,dt_1 \ldots dt_n.$$

Din această formulă de inversiune se deduce totodată că funcția de repartiție $F(x_1, \ldots, x_n)$ este unic determinată de funcția caracteristică $\varphi(t_1, \ldots, t_n)$.

Scanned with OKEN Scanner

8.4. LEGĂTURA DINTRE FUNCȚIA CARACTERISTICĂ ȘI PRODUSUL DE COMPOZIȚIE

Între produsul de compoziție al funcțiilor de repartiție (vezi § 5.1) și produsul funcțiilor caracteristice există o legătură remarcabilă, care este precizată în teorema următoare.

Teorema 8.4.1. *Fie* $F_0, F_1, \ldots, F_n$ *funcții de repartiție și fie* $\varphi_0, \varphi_1, \ldots, \varphi_n$ *funcțiile caracteristice corespunzătoare. Condiția necesară și suficientă ca* F_0 *să fie produsul de compoziție al funcțiilor* $F_1, \ldots, F_n$, *adică*

$$F_0 = F_1 * \ldots * F_n \tag{8.4.1}$$

este ca

$$\varphi_0 = \varphi_1 \ldots \varphi_n. \tag{8.4.2}$$

Demonstrație. Vom face demonstrația pentru cazul $n = 2$, cazul general obținîndu-se evident prin inducție.

Pentru a arăta necesitatea condiției, să considerăm că $F_0 = F_1 * F_2$. Ne fixăm atenția asupra unui interval $[a, b] \in R$ și considerăm un șir de diviziuni

$$a = x_0^{(n)} < x_1^{(n)} < \ldots < x_{k(n)}^{(n)} = b$$

ale acestui interval, cu proprietatea că

$$\lim_n \max_{1 \leqslant j \leqslant k(n)} (x_j^{(n)} - x_{j-1}^{(n)}) = 0.$$

Atunci, conform definiției integralei Stieltjes, putem scrie

$$\varphi_0(t) = \int_a^b e^{itx}\, dF_0(x) = \lim_{n \to \infty} \sum_{j=1}^{k(n)} e^{it x_{j-1}^{(n)}} [F_0(x_j^{(n)}) - F_0(x_{j-1}^{(n)}].$$

Deoarece $F_0 = F_1 * F_2$, adică

$$F_0(x) = \int_R F_1(x-y)\, dF_2(y),$$

rezultă

$$\varphi_0(t) = \lim_{n\to\infty} \int_R \sum_{j=1}^{k(n)} e^{it(x_{j-1}^{(n)} - y)} [F_1(x_j^{(n)} - y) - F_1(x^{(n)} - y)]\, e^{ity}\, dF_2(y) =$$

$$= \int_R \left(\int_{a-y}^{a+y} e^{itx}\, dF_1(x) \right) e^{ity}\, dF_2(y).$$

Trecînd la limită pentru $a \to -\infty$ și $b \to +\infty$, obținem

$$\varphi_0(t) = \left(\int_R e^{itx}\, dF_1(x) \right) \left(\int_R e^{ity}\, dF_2(y) \right)$$

sau

$$\varphi_0(t) = \varphi_1(t)\, \varphi_2(t),$$

ceea ce trebuia demonstrat.

Pentru demonstrarea suficienței, să considerăm $\varphi = \varphi_1\varphi_2$. Din prima parte a teoremei rezultă că $\varphi_1\varphi_2$ este funcția caracteristică a lui $F_1 * F_2$. Însă funcția de repartiție este unic determinată de funcția caracteristică, de unde rezultă că $F = F_1 * F_2$, ceea ce încheie demonstrația.

Observații. Această teoremă furnizează o justificare a afirmației că funcția de repartiție a unei sume de variabile aleatoare independente este produsul de compoziție al funcțiilor de repartiție ale termenilor, dacă ținem seama și de proprietatea (h) din teorema 8.2.1.

Rezultă de asemenea o justificare pentru faptul că produsul de compoziție este comutativ și asociativ.

Un produs de funcții caracteristice este tot o funcție caracteristică. Ținînd seama că $\overline{\varphi}_{\tilde{x}}$ este funcția caracteristică a lui $-\tilde{x}$, rezultă că

$$|\varphi|^2 = \varphi_{\tilde{x}} \cdot \overline{\varphi}$$

este de asemenea o funcție caracteristică.

8.5. TEOREME DE CONVERGENȚĂ

În cele ce urmează vom stabili legătura dintre convergența funcțiilor de repartiție și convergența funcțiilor caracteristice corespunzătoare.

Fie $(\widetilde{x}_n)_{n\geqslant 0}$ un șir de variabile aleatoare pe cîmpul borelian de probabilitate $(E, \mathfrak{K}, P)$ și fie $(F_n)_{n\geqslant 0}$ și $(\varphi_n)_{n\geqslant 0}$ respectiv șirul funcțiilor de repartiție și șirul funcțiilor caracteristice corespunzătoare acestui șir de variabile aleatoare.

Teorema 8.5.1. *Dacă șirul* $(F_n)_{n\geqslant 1}$ *converge în repartiție spre* F_0, *atunci:*

(a) Șirul $(\varphi_n)_{n\geqslant 1}$ *converge spre* φ_0 *în orice punct* $t\in R$;

(b) Șirul $(\varphi_n)_{n\geqslant 1}$ *converge uniform spre* φ_0 *pe orice interval* $[a, b]\subset R$.

Demonstrație. (a). Deoarece e^{itx} este continuă și mărginită pe R, rezultatul se deduce imediat din definiția convergenței în repartiție a funcțiilor de repartiție.

(b). Demonstrația acestei afirmații este un exercițiu de analiză matematică analog celui utilizat în demonstrația teoremei (7.4.1) și o lăsăm în seama cititorului.

Pentru a obține legătura inversă dintre convergența șirurilor (F_n) și (φ_n), vom prezenta mai întîi un rezultat de analiză matematică cunoscut sub numele de *prima teoremă a lui Helly*.

Teorema 8.5.2. *Dacă* (F_n) *este un șir de funcții crescătoare și uniform mărginite pe* R, *atunci există un subșir* (F_{n_k}) *al acestui șir care converge către o funcție crescătoare* F_0 *în toate punctele de continuitate ale acesteia din urmă.*

Demonstrație. Fie $A\subset R$ o mulțime numărabilă, densă în R (astfel de mulțimi există, de exemplu mulțimea numerelor raționale). Deoarece A este numărabilă, putem scrie $A = (a_n)_{n\in N}$.

Șirul $(F_n(a_1))$ de numere reale este mărginit și deci putem extrage din el un subșir convergent $(F_n^{(1)}(a_1))$. Analog, $(F_n^{(1)}(a_2))$ este un șir mărginit și deci putem determina un subșir convergent $(F_n^{(2)}(a_2))$ al acestui șir. Evident, $(F_n^{(2)}(a_1))$ este de asemenea convergent. În general, la etapa k, determinăm un

subșir $(F_n^{(k)})$ al lui (F_n) cu proprietatea că $(F_n^{(k)}(a_i))$ cu $1 \leqslant i \leqslant k$, sînt șiruri convergente de numere reale. Vom nota

$$\lim_n F_n^{(k)}(a_i) = F_A(a_i), \quad 1 \leqslant i \leqslant k.$$

Să considerăm acum șirul diagonal $(F_n^{(n)})_{n \in N}$. Este clar că șirul $(F_n^{(n)})$, cu excepția primilor $k-1$ termeni, este un subșir al lui $(F_n^{(k)})$ și deci

$$\lim_n F_n^{(n)}(a_i) = F_A(a_i) \tag{8.5.1}$$

pentru orice $a_i \in A$. Cu alte cuvinte, notînd (F_{n_k}) șirul diagonal $(F_n^{(n)})$, din relația (8.5.1) rezultă că putem defini funcția $F_A : A \to R$ prin

$$F_A(x) = \lim_k F_{n_k}(x). \tag{8.5.2}$$

Este clar că funcția F_A astfel definită este crescătoare și mărginită pe A. Fie acum

$$F_0(x) = \sup_{\substack{y<x \\ y \in A}} F_A(y). \tag{8.5.3}$$

Se vede ușor că F_0 este o funcție crescătoare și mărginită pe R.

Fie acum $x \in R$, $x' \leqslant x \leqslant x''$ cu x', $x'' \in A$. Avem evident

$$F_{n_k}(x') \leqslant F_{n_k}(x) \leqslant F_{n_k}(x'').$$

Trecînd la limită obținem

$$\lim_k F_{n_k}(x') \leqslant \underline{\lim_k} F_{n_k}(x) \leqslant \overline{\lim_k} F_{n_k}(x) \leqslant \lim_k F_{n_k}(x'')$$

sau

$$F_0(x') \leqslant \underline{\lim_k} F_{n_k}(x) \leqslant \overline{\lim_k} F_{n_k}(x) \leqslant F_0(x''). \tag{8.5.4}$$

Deoarece A este densă în R, trecînd la limită în (8.5.4) pentru $x' \to x$ și $x'' \to x$ cu x', $x'' \in A$, obținem

$$F_0(x-0) \leqslant \underline{\lim_k} F_{n_k}(x) \leqslant \overline{\lim_k} F_{n_k}(x) \leqslant F_0(x+0). \tag{8.5.5}$$

Scanned with OKEN Scanner

Dacă x este punct de continuitate al lui F_0, din (8.5.5) rezultă

$$\lim_k F_{n_k}(x) = F_0(x)$$

și astfel teorema este demonstrată.

Observație. Dacă (F_n) este un șir de funcții de repartiție rezultă că îndeplinește condițiile teoremei și deci există un subșir (F_{n_k}) și o funcție F_0 crescătoare și cu valori în $[0, 1]$ cu proprietatea că (F_{n_k}) converge spre F_0 în toate punctele de continuitate ale lui F_0. Evident, este posibil ca F_0 să nu fie continuă la stînga. Dacă însă punem

$$F_0^*(x) = \begin{cases} F_0(x) & \text{dacă } x \text{ este punct de continuitate al lui } F_0, \\ F_0(x-0) & \text{în caz contrar,} \end{cases}$$

rezultă că F_0^* este crescătoare, are valori în $[0, 1]$, este continuă la stînga și (F_{n_k}) converge spre F_0^* în toate punctele de continuitate ale lui F_0^*. Se poate arăta însă că F_0^* nu este în mod necesar o funcție de repartiție (vezi exemplul dat la teorema următoare).

Teorema 8.5.3. *Dacă șirul de funcții caracteristice* $(\varphi_n(t))$ *converge pentru orice* $t \in R$ *către o funcție* $\varphi_0(t)$ *continuă în* $t=0$, *atunci șirul corespunzător de funcții de repartiție* (F_n) *converge în repartiție către o funcție de repartiție* F_0. *În plus,* φ_0 *este funcția caracteristică corespunzătoare lui* F_0.

Demonstrație. Din teorema precedentă (vezi și observația de mai sus) rezultă că din șirul (F_n) putem extrage un subșir (F_{n_k}) convergent către o funcție F_0 crescătoare, cu valori în $[0, 1]$ și continuă la stînga, convergența avînd loc în toate punctele de continuitate ale lui F_0.

Vom arăta că, în condițiile teoremei noastre, F_0 este chiar o funcție de repartiție, adică $F_0(-\infty) = 0$ și $F_0(+\infty) = 1$. Să presupunem prin absurd că

$$a = F_0(+\infty) - F_0(-\infty) < 1. \tag{8.5.6}$$

Deoarece $(\varphi_n(0))$ converge spre $\varphi_0(0)$, rezultă că $\varphi_0(0) = 1$.

Scanned with OKEN Scanner

Deoarece φ_0 este continuă în $t = 0$, rezultă că pentru orice $\varepsilon > 0$ putem determina un număr $\tau > 0$ suficient de mic astfel încît

$$\frac{1}{2\tau}\left|\int_{-\tau}^{\tau}\varphi_0(t)\,\mathrm{d}t\right| > 1 - \varepsilon/2.$$

Dacă alegem $\varepsilon < 1 - \delta$, rezultă că

$$\frac{1}{2\tau}\left|\int_{-\tau}^{\tau}\varphi_0(t)\,\mathrm{d}t\right| > \delta + \varepsilon/2. \tag{8.5.7}$$

Pentru orice număr real $M > 0$ avem însă

$$\left|\int_{-\tau}^{\tau}\varphi_{n_k}(t)\,\mathrm{d}t\right| \leqslant \left|\int_{|x|\leqslant M}\left[\int_{-\tau}^{\tau}\mathrm{e}^{\mathrm{i}tx}\,\mathrm{d}t\right]\mathrm{d}F_{n_k}(x)\right| +$$

$$+\left|\int_{|x|>M}\left[\int_{-\tau}^{\tau}\mathrm{e}^{\mathrm{i}tx}\,\mathrm{d}t\right]\mathrm{d}F_{n_k}(x)\right| \leqslant \int_{|x|\leqslant M} 2\tau\,\mathrm{d}F_{n_k}(x) +$$

$$+\int_{|x|>M}\frac{2|\sin\tau x|}{|x|}\,\mathrm{d}F_{n_k}(x) \leqslant 2\tau\int_{|x|\leqslant M}\mathrm{d}F_{n_k}(x) + (2/M)\int_{|x|>M}\mathrm{d}F_{n_k}(x)$$

și deci

$$(1/2\,\tau)\left|\int_{-\tau}^{\tau}\varphi_{n_k}(t)\,\mathrm{d}t\right| \leqslant \int_{|x|\leqslant M}\mathrm{d}F_{n_k}(x) + (1/M\,\tau). \tag{8.5.8}$$

Să alegem acum $M > 4/\tau\varepsilon$. Din (8.5.6) rezultă că există $K(\varepsilon)$ cu proprietatea

$$\int_{|x|\leqslant M}\mathrm{d}F_{n_k}(x) < \delta + \varepsilon/4 \tag{8.5.9}$$

Scanned with OKEN Scanner

pentru orice $k \geqslant K(\varepsilon)$. Din (8.5.8) și (8.5.9) obținem atunci

$$(1/2\,\tau)\left|\int_{-\tau}^{\tau} \varphi_{n_k}(t)\,\mathrm{d}t\right| < \delta+\varepsilon/2,$$

ceea ce contrazice (8.5.7).

Rezultă deci că (F_{n_k}) converge în repartiție spre F_0 și deci

$$\lim_k \int_R \mathrm{e}^{\mathrm{i}tx}\,\mathrm{d}F_{n_k}(x) = \int_R \mathrm{e}^{\mathrm{i}tx}\,\mathrm{d}F_0(x) = \varphi_0(t).$$

Pentru a termina demonstrația trebuie să mai arătăm că șirul (F_n) converge în repartiție spre F_0. Într-adevăr, să presupunem prin absurd că există un subșir (F_{m_k}) al lui (F_n) convergent în repartiție spre o funcție de repartiție F și $F \neq F_0$. Să observăm însă că funcția caracteristică a lui F este tot φ_0 și din teorema 8.3.1 rezultă că trebuie să avem $F = F_0$.

Observație. Teorema oferă evident posibilitatea de a demonstra convergența în repartiție a șirului de funcții de repartiție (F_n) către o funcție de repartiție dată F_0. Importanța teoremei constă mai ales în posibilitatea de a demonstra că un șir de funcții de repartiție dat (F_n) este convergent în repartiție, putînd astfel defini noi funcții de repartiție.

Se poate arăta printr-un exemplu că ipoteza ca φ_0 să fie continuă în $t = 0$ este esențială pentru valabilitatea teoremei. Într-adevăr, fie (F_n) șirul de funcții de repartiție definite de densitățile de probabilitate

$$f_n(x) = \begin{cases} 1/2n & \text{pentru } x \in [-n, n], \\ 0 & \text{pentru } x \in (-\infty, -n) \cup (n, +\infty). \end{cases}$$

Funcțiile caracteristice corespunzătoare se calculează ușor și sînt date de

$$\varphi_n(t) = \begin{cases} \dfrac{\sin(nt)}{nt} & \text{pentru } t \neq 0, \\ 1 & \text{pentru } t = 0. \end{cases}$$

Scanned with OKEN Scanner

Este clar că șirul $(\varphi_n(t))$ este convergent în fiecare punct $t \in R$ către o funcție $\varphi(t)$ dată de relația

$$\varphi(t) = \begin{cases} 0 \text{ pentru } t \neq 0, \\ 1 \text{ pentru } t = 0, \end{cases}$$

care nu este continuă în $t = 0$.

Șirul de funcții de repartiție (F_n) nu converge către o funcție de repartiție, deoarece

$$F_n(x) = \begin{cases} 0 & \text{pentru } x \in (-\infty, -n], \\ \dfrac{x+n}{2n} & \text{pentru } x \in (-n, n], \\ 1 & \text{pentru } x \in (n, +\infty) \end{cases}$$

și deci

$$\lim_n F_n(x) = 1/2.$$

Amintim că teoreme de convergență analoage au loc și în cazul multidimensional.

8.6. FUNCȚII GENERATOARE

Să considerăm o funcție $g : N \to R$ cu proprietatea că există un număr real $a > 0$ astfel încît

$$|g(n)| \leqslant a^n \tag{8.6.1}$$

pentru orice $n \in N$. Vom nota prin G_a mulțimea funcțiilor care satisfac condiția (8.6.1).

Definiția 8.6.1. *Funcția* g^+ *de variabilă complexă* z *definită de relația*

$$g^+(z) = \sum_{n \in N} g(n) z^n \tag{8.6.2}$$

este numită funcția generatoare a funcției $g \in G_a$.

Este clar că seria din membrul drept al relației (8.6.2) este uniform convergentă pe domeniul D_a din planul complex definit de

$$D_a = \{z : |z| < 1/a\} \tag{8.6.3}$$

și deci g^+ este o funcție olomorfă în D_a. Vom nota prin G_a^+ mulțimea funcțiilor olomorfe definite pe D_a.

Definiția 8.6.2. *Vom numi transformarea z aplicația* $G_a \to G_a^+$ *care face să corespundă fiecărui element* $g \in G_a$ *funcția sa generatoare* $g^+ \in G_a^+$.

Este ușor de văzut că aplicația astfel definită este bijectivă. Într-adevăr, orice $g^+ \in G_a^+$ poate fi dezvoltată *în mod unic* în serie de puteri în domeniul D_a și deci coeficienții acestei dezvoltări în serie definesc în mod unic o funcție $g \in G_a$. Funcția g^+ se mai numește și *transformata în z* a funcției g.

Tabelul 8.6.1 conține principalele proprietăți ale transformării z, a căror demonstrație nu comportă nici o dificultate.

Tabelul 8.6.1

Funcția	Transformata în z
$g(n) = g_1(n) + g_2(n)$	$g^+(z) = g_1^+(z) + g_2^+(z)$
$G(n) = ag(n),\ a \in R$	$G^+(z) = ag^+(z)$
$G(n) = \begin{cases} g(n-1), & n > 1 \\ 0, & n = 0 \end{cases}$	$G^+(z) = zg^+(z)$
$G(n) = g(n+1)$	$G^+(z) = (g^+(z) - g(0))/z$
$G(n) = a^n g(n),\ \lvert a\rvert \leqslant 1$	$G^+(z) = g^+(az)$
$G(n) = \sum_{i=0}^{n} g(i)$	$G^+(z) = g^+(z)/(1-z)$

Mai adăugăm la acestea faptul că

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} g(i) = \lim_{z\to 1} (1-z) g^+(z) \tag{8.6.4}$$

ori de cîte ori una dintre cele două limite există (pentru demonstrație vezi, de exemplu, Hardy, G.H. (1949)). Evident, dacă $(g(n))$ este un șir convergent, atunci avem chiar

$$\lim_n g(n) = \lim_{z\to 1} (1-z)g^+(z). \tag{8.6.5}$$

Principalele transformate în z sînt prezentate în tabelul 8.6.2.

Tabelul 8.6.2

Funcția	Transformata în z
$g(n) = 1$	$g^+(z) = 1/(1-z)$
$g(n) = n$	$g^+(z) = z/(1-z)^2$
$g(n) = an,\ a\in R$	$g^+(z) = az/(1-z)^2$
$g(n) = a^n,\ \|a\|\leqslant 1$	$g^+(z) = 1/(1-az)$
$g(n) = na^n,\ \|a\|\leqslant 1$	$g^+(z) = az/(1-az)^2$
$g(n) = C^n_{n+p-1}a^n,\ \|a\|\leqslant 1$	$g^+(z) = 1/(1-az)^p$
$g(n) = C^n_{n+p-2}a^n,\ \|a\|\leqslant 1$	$g^+(z) = az/(1-az)^p$

Dacă $\mathbf{G}(n) = (g_{ij}(n))$ este o matrice ale cărei elemente sînt funcții cu valori reale care satisfac condiția (8.6.1), atunci matricea $\mathbf{G}^+(z) = (g^+_{ij}(z))$ ale cărei elemente sînt transformatele în z ale funcțiilor $g_{ij}(n)$ este de asemenea numită *transformata în z* sau *matricea generatoare* a matricei $\mathbf{G}(n)$. Este ușor de văzut că proprietățile transformării z prezentate mai sus se extind fără dificultate la cazul matriceal.

Să considerăm o variabilă aleatoare discretă $\tilde{x}$ definită pe cîmpul borelian de probabilitate $(E, \mathfrak{K}, P)$. Fără a restrînge generalitatea, ne putem mărgini la considerarea cazului cînd mulțimea valorilor posibile ale lui $\tilde{x}$ este inclusă în mulțimea numerelor naturale N. Putem atunci introduce funcția $p_{\tilde{x}} : N\to R$ definită de

$$p_{\tilde{x}}(n) = P(\tilde{x} = n),\quad n\in N. \tag{8.6.6}$$

Este ușor de văzut că $p_{\tilde{x}}$ este restricția la N a funcției de masă f a variabilei aleatoare $\tilde{x}$; avem deci $p_{\tilde{x}}(n) = f(n)$.

Este clar că funcția $p_{\tilde{x}}$ astfel definită verifică condiția (8.6.1), deoarece $|p_{\tilde{x}}(n)| = p_{\tilde{x}}(n) \leqslant 1$ pentru orice $n \in N$. Putem deci introduce funcția generatoare

$$p_{\tilde{x}}^{+}(z) = \sum_{n \in N} p_{\tilde{x}}(n) z^n \tag{8.6.7}$$

definită pentru $|z| \leqslant 1$.

Definiția 8.6.3. *Funcția* $p_{\tilde{x}}^{+}(z)$ *definită de* (8.6.7) *este numită funcția generatoare a variabilei aleatoare discrete* $\tilde{x}$ *avînd mulțimea valorilor* N.

Cînd nu este posibilă o confuzie, vom scrie $p^{+}(z)$ în loc de $p_{\tilde{x}}^{+}(z)$.

Funcția caracteristică a variabilei aleatoare $\tilde{x}$ este

$$\varphi_{\tilde{x}}(t) = M(e^{it\tilde{x}}) = \sum_{n \in N} e^{itn} p_{\tilde{x}}(n) = p_{\tilde{x}}^{+}(e^{it}), \tag{8.6.8}$$

ceea ce arată strînsa legătură dintre funcția caracteristică și funcția generatoare asociate unei variabile aleatoare discrete.

Pentru funcția generatoare a unei variabile-aleatoare discrete mai menționăm și proprietățile precizate în teorema următoare.

Teorema 8.6.1. *Fie* $\tilde{x}$ *și* $\tilde{y}$ *variabile aleatoare discrete pe cîmpul borelian de probabilitate* (E, $\mathfrak{K}$, P), *avînd mulțimea valorilor posibile* N.

(a) Avem

$$\lim_{n} (1/n) \sum_{i=0}^{n=1} p_{\tilde{x}}(i) = \lim_{z \to 1} (1-z)\, p_{\tilde{x}}^{+}(z); \tag{8.6.9}$$

(b) Dacă $q_{\tilde{x}} : N \to R$ *este definită prin*

$$q_{\tilde{x}}(n) = \sum_{k=n+1}^{\infty} p_{\tilde{x}}(k), \tag{8.6.10}$$

atunci

$$q_{\tilde{x}}^{+}(z) = \frac{1 - p_{\tilde{x}}^{+}(z)}{1 - z}; \tag{8.6.11}$$

(c) Dacă $\tilde{x} \in \mathcal{L}(E, \mathfrak{K}, P)$, *atunci*

$$M(\tilde{x}) = q_{\tilde{x}}^{+}(1) = \frac{d}{dz}\,(p_{\tilde{x}}^{+}(z))_{z=1}; \tag{8.6.12}$$

(d) Dacă $\tilde{x}$ *și* $\tilde{y}$ *sînt variabilele aleatoare independente, atunci*

$$p_{\tilde{x}+\tilde{y}}^{+}(z) = p_{\tilde{x}}^{+}(z) \cdot p_{\tilde{y}}^{+}(z). \tag{8.6.13}$$

Demonstrație. (a). Este evident că limita din membrul stîng al relației (8.6.9) există, deoarece

$$0 \leqslant \sum_{i=0}^{n-1} p_{\tilde{x}}(i) \leqslant 1$$

pentru orice $n \geqslant 1$. Afirmația decurge atunci din proprietatea (8.6.4) a transformării z.

(b). Putem scrie

$$q_{\tilde{x}}(n) = 1 - \sum_{k=0}^{n} p_{\tilde{x}}(k)$$

și deci

$$q_{\tilde{x}}^{+}(z) = \sum_{n=0}^{\infty}\left(1 - \sum_{k=0}^{n} p_{\tilde{x}}(k)\right)z^n = 1/(1-z) - p_{\tilde{x}}^{+}(z)/(1-z),$$

dacă ținem seama de proprietățile transformării z prezentate în tabelul 8.6.1.

(c). Avem

$$M(\tilde{x}) = \sum_{n=0}^{\infty} n p_{\tilde{x}}(n) = \sum_{n=0} q_{\tilde{x}}(n) = q_{\tilde{x}}^{+}(1)$$

și prima egalitate din (8.6.12) este astfel obținută.

Deoarece

$$\frac{d}{dz}\,(p_{\tilde{x}}^{+}(z)) = \frac{d}{dz}\left(\sum_{n=0}^{\infty} p_{\tilde{x}}(n)z^n\right) = \sum_{n=1}^{\infty} n p_{\tilde{x}}(n) z^{n-1} =$$

$$= (1/z)\sum_{n=0}^{\infty} n p_{\tilde{x}}(n) z^n,$$

a doua egalitate din (8.6.12) se deduce imediat.

Scanned with OKEN Scanner

(d) Avem

$$p_{\tilde{x}+\tilde{y}}(n) = P(\tilde{x}+\tilde{y}=n) = \sum_{k=0}^{n} P(\tilde{x}=k)\, P(\tilde{x}+\tilde{y}=$$

$$= n \mid \tilde{x}=k) = \sum_{k=0}^{n} P(\tilde{x}=k)\, P(\tilde{y}=n-k \mid \tilde{x}=k). \quad (8.6.14)$$

Ținînd seama că $\tilde{x}$ și $\tilde{y}$ sînt independente, obținem

$$P(\tilde{y}=n-k \mid \tilde{x}=k) = P(\tilde{y}=n-k)$$

și deci relația (8.6.14) devine

$$p_{\tilde{x}+\tilde{y}}(n) = \sum_{k=0}^{n} p_{\tilde{x}}(k)\, p_{\tilde{y}}(n-k).$$

Rezultă atunci

$$p^{+}_{\tilde{x}+\tilde{y}}(z) = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} p_{\tilde{x}}(k)\, p_{\tilde{y}}(n-k)\right)z^n =$$

$$= \sum_{n=0}^{\infty} p_{\tilde{x}}(n)\, p^{+}_{\tilde{y}}(z) z^n = p^{+}_{\tilde{x}}(z)\, p^{+}_{\tilde{y}}(z),$$

ceea ce demonstrează afirmația.

Observație. Dacă $\tilde{x}$ are momente de ordin superior, atunci acestea se pot calcula prin formule analoage formulei (8.6.12). De exemplu, dacă $\tilde{x}$ are moment de ordinul al doilea, atunci

$$M_2(\tilde{x}) = \sum_{n=0}^{\infty} n^2 p_{\tilde{x}}(n) = \sum_{n=0}^{\infty} (n^2-n)\, p_{\tilde{x}}(n) + \sum_{n=0}^{\infty} n p_{\tilde{x}}(n) = \quad (8.6.15)$$

$$= \frac{\mathrm{d}^2}{\mathrm{d}z^2}\left(p^{+}_{\tilde{x}}(z)\right)\Big|_{z=1} + \frac{\mathrm{d}}{\mathrm{d}z}\left(p^{+}_{\tilde{x}}(z)\right)\Big|_{z=1}.$$

Proprietatea *(d)* poate fi generalizată imediat. Într-adevăr, dacă $\tilde{x}_1, \ldots, \tilde{x}_m$ sînt variabile aleatoare independente discrete, atunci

$$p^{+}_{\tilde{x}_1+\ldots+\tilde{x}_m}(z) = p^{+}_{\tilde{x}_1}(z) \ldots p^{+}_{\tilde{x}_m}(z). \quad (8.6.16)$$

În sfîrșit, se poate observa că proprietatea (8.6.13) nu este suficientă pentru ca $\tilde{x}$ și $\tilde{y}$ să fie independente.

Vom încheia considerațiile privind funcția generatoare cu teoreme referitoare la convergența șirurilor de funcții generatoare, analoage teoremelor 8.5.1 și 8.5.3.

Teorema 8.6.2. *Fie* $(\tilde{x}_m)_{m\geqslant 0}$ *un șir de variabile aleatoare discrete pe cîmpul borelian de probabilitate* $(E, \mathcal{K}, P)$ *și avînd mulțimea valorilor* N. *Dacă*

$$\lim_m p_{\tilde{x}_m}(n) = p_{\tilde{x}_0}(n) \tag{8.6.17}$$

pentru fiecare $n \in N$, *atunci*

$$\lim_m p^{\pm}_{\tilde{x}_m}(z) = p^{\pm}_{\tilde{x}_0}(z) \tag{8.6.18}$$

pentru $|z| \leqslant 1$.

Demonstrație. Să considerăm mai întîi $|z| < 1$ și fie $K \geqslant 1$ un număr natural arbitrar. Putem atunci scrie

$$|p^{\pm}_{\tilde{x}_m}(z) - p^{\pm}_{\tilde{x}_0}(z)| = \left|\sum_{k=0}^{\infty} (p_{\tilde{x}_m}(k) - p_{\tilde{x}_0}(k)) z^k\right| \leqslant$$
$$\leqslant \sum_{k=0}^{K-1} |p_{\tilde{x}_m}(k) - p_{\tilde{x}_0}(k)| + |z|^K \sum_{k=K}^{\infty} [p_{\tilde{x}_m}(k) - p_{\tilde{x}_0}(k)]. \tag{8.6.19}$$

Pentru orice $\varepsilon > 0$, deoarece $|z| < 1$, putem alege K suficient de mare astfel încît

$$|z|^K < \varepsilon/4 \tag{8.6.20}$$

și apoi, deoarece avem (8.6.17), putem alege $M_1(\varepsilon)$ suficient de mare astfel încît

$$|p_{\tilde{x}_m}(k) - p_{\tilde{x}_0}(k)| < \varepsilon/2K \tag{8.6.21}$$

pentru toți $k = 0, 1, \ldots, K-1$ și orice $m \geqslant M_1(\varepsilon)$.

Din (8.6.19), (8.6.20) și (8.6.21) obținem

$$|p^{\pm}_{\tilde{x}_m}(z) - p^{\pm}_{\tilde{x}_0}(z)| < K\,\varepsilon/2K + 2\varepsilon/4 = \varepsilon$$

pentru $m \geqslant M_1(\varepsilon)$, de unde se deduce (8.6.18) în acest caz.

Scanned with OKEN Scanner

Fie acum cazul $|z| = 1$. Deoarece seria $\sum_{k=0}^{\infty} p_{x_0}(k)$ este convergentă, pentru $\varepsilon > 0$ dat putem determina K suficient de mare astfel încît

$$\sum_{k=K}^{\infty} p_{x_0}^{\sim}(k) < \varepsilon/4. \tag{8.6.22}$$

Să observăm că avem

$$\lim_m \sum_{k=K}^{\infty} p_{x_m}^{\sim}(k) = \lim_m \left(1 - \sum_{k=0}^{K-1} p_{x_m}^{\sim}(k)\right) =$$

$$= 1 - \sum_{k=0}^{K-1} p_{x_0}^{\sim}(k) = \sum_{k=K}^{\infty} p_{x_0}^{\sim}(k)$$

și deci putem determina $M_2(\varepsilon)$ suficient de mare pentru a obține

$$\left| \sum_{k=K}^{\infty} p_{x_m}^{\sim}(k) - \sum_{k=K}^{\infty} p_{x_0}^{\sim}(k) \right| < \varepsilon/4 \tag{8.6.23}$$

pentru orice $m \geqslant M_2(\varepsilon)$.

Din relațiile (8.6.22) și (8.6.23) obținem atunci

$$\sum_{k=K}^{\infty} p_{x_m}^{\sim}(k) \leqslant \left| \sum_{k=K}^{\infty} [p_{x_m}^{\sim}(k) - p_{x_0}^{\sim}(k)] \right| + \sum_{k=K}^{\infty} p_{x_0}^{\sim}(k) < \varepsilon/2 \tag{8.6.24}$$

pentru $m \geqslant M_2(\varepsilon)$.

Utilizînd din nou (8.6.17), putem alege $M_3(\varepsilon)$ suficient de mare astfel încît

$$| p_{x_m}^{\sim}(k) - p_{x_0}^{\sim}(k) | < \varepsilon/4K \tag{8.6.25}$$

pentru toți $k = 0, 1, \ldots, K-1$ și orice $m \geqslant M_3(\varepsilon)$.

Din relațiile (8.6.19), (8.6.22), (8.6.24) și (8.6.25) obținem atunci

$$| p_{x_m}^{\pm}(z) - p_{x_0}^{\pm}(z) | < \varepsilon K/4K + \varepsilon/2 + \varepsilon/4 = \varepsilon$$

pentru toți $m \geqslant M(\varepsilon) = \max(M_2(\varepsilon), M_3(\varepsilon))$, de unde rezultă relația (8.6.18) în acest caz.

Scanned with OKEN Scanner

Observație. Dacă există limitele

$$\lim_m p_{\tilde{x}_m}(n) = p(n)$$

pentru fiecare $n \in N$, vom avea, evident, $0 \leqslant p(n) \leqslant 1$, dar, în general, $\sum_{n=0}^{\infty} p(n)$ nu mai este în mod necesar convergentă cu suma egală cu 1. Concluzia (8.6.18) se menține însă și în acest caz, dar numai pentru $|z| < 1$, cum rezultă din prima parte a demonstrației dată mai sus.

Teorema precedentă poate fi enunțată în mod echivalent astfel: dacă $(\tilde{x}_m)$ converge în repartiție spre $\tilde{x}_0$, atunci $(p^{\pm}_{\tilde{x}_m}(z))$ converge spre $p^{\pm}_{\tilde{x}_0}(z)$ pentru $|z| \leqslant 1$.

Teorema 8.6.3. *Fie* $(\tilde{x}_n)$ *un șir de variabile aleatoare discrete având valori întregi nenegative și* $(p^{\pm}_{\tilde{x}_n})$ *șirul funcțiilor generatoare corespunzătoare. Dacă șirul* $(p^{\pm}_{\tilde{x}_n})$ *converge pentru orice* $|z| \leqslant 1$ *către o funcție generatoare*

$$p^+(z) = \sum_{k \in N} p(k) z^k, \tag{8.6.26}$$

unde $0 \leqslant p(k) \leqslant 1$, $k \in N$, *atunci*

$$\lim_n p_{\tilde{x}_n}(k) = p(k) \tag{8.6.27}$$

pentru orice $k \in N$ *și, în plus, avem*

$$\sum_{k \in N} p(k) = 1, \tag{8.6.28}$$

adică $p^+(z)$ *este funcția generatoare a unei variabile aleatoare* $\tilde{x}_0$ *cu valori întregi nenegative pentru care*

$$p_{\tilde{x}_0}(k) = p(k), \quad k \in N.$$

Demonstrație. Afirmația (8.6.27) o demonstrăm prin inducție după $k \in N$. Pentru $k = 0$, avem evident

$$\lim_n p_{\tilde{x}_n}(0) = \lim_n p^{\pm}_{\tilde{x}_n}(0) = p^+(0) = p(0).$$

Să presupunem că (8.6.27) este adevărată pentru $0 \leqslant k \leqslant m-1$ și să arătăm că este adevărată și pentru $k = m$. Să observăm mai întîi că

$$\frac{\mathrm{d}^m}{\mathrm{d}z^m}\left(p^{\pm}_{\tilde{x}_n}(z)\right) = z^{-m}\left(p^{\pm}_{\tilde{x}_n}(z) - \sum_{k=0}^{m-1} p_{\tilde{x}_n}(k)\, z^k\right) \tag{8.6.29}$$

și, analog,

$$\frac{\mathrm{d}^m}{\mathrm{d}z^m}\left(p^{+}(z)\right) = z^{-m}\left(p^{+}(z) - \sum_{k=0}^{m-1} p(k) z^k\right). \tag{8.6.30}$$

Din (8.6.29) și (8.6.30) rezultă

$$\lim_m \frac{\mathrm{d}^m}{\mathrm{d}z^m}\left(p^{\pm}_{\tilde{x}_n}(z) - p^{+}(z)\right) = 0 \tag{8.6.31}$$

pentru $|z| \leqslant 1$.

Utilizînd (8.6.29) și (8.6.30) putem scrie

$$p_{\tilde{x}_n}(m) - p(m) = \frac{\mathrm{d}^m}{\mathrm{d}z^m}\left(p^{\pm}_{\tilde{x}_n}(z) - p^{+}(z)\right) - \sum_{k=m+1}^{\infty} \left(p_{\tilde{x}_n}(k) - p(k)\right) z^{k-m}$$

și deci

$$|p_{\tilde{x}_n}(m) - p(m)| \leqslant \left|\frac{\mathrm{d}^m}{\mathrm{d}z^m}\left[p^{\pm}_{\tilde{x}_n}(z) - p^{+}(z)\right]\right| + |z|/(1-|z|), \tag{8.6.32}$$

dacă se ține seama că

$$|p_{\tilde{x}_n}(m) - p(m)| \leqslant 1.$$

Fie $\varepsilon > 0$ și să alegem z_0 astfel încît $|z_0| < \varepsilon/(2+\varepsilon)$ și apoi să alegem $N(\varepsilon)$ astfel încît

$$\left|\frac{\mathrm{d}^m}{\mathrm{d}z^m}\left(p^{\pm}_{\tilde{x}_n}(z_0) - p^{+}(z_0)\right)\right| < \varepsilon/2 \tag{8.6.33}$$

pentru orice $n \geqslant N(\varepsilon)$.

Din relațiile (8.6.32) și (8.6.33) obținem atunci

$$|p_{\tilde{x}_n}(m) - p(m)| < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

Scanned with OKEN Scanner

pentru orice $n \geqslant N(\varepsilon)$, ceea ce arată că

$$\lim_n p_{\tilde{x}_n}(m) = p(m)$$

și astfel demonstrația prin inducție a afirmației (8.6.27) este încheiată.

Pentru a arăta și (8.6.28), să observăm că pentru $z = 1$, din ipoteză rezultă că

$$\lim_n p^+_{\tilde{x}_n}(1) = p^+(1). \tag{8.6.34}$$

Avem însă, evident,

$$p^+_{\tilde{x}_n}(1) = \sum_{k=0}^{\infty} p_{\tilde{x}_n}(k) = 1 \tag{8.6.35}$$

și concluzia (8.6.28) se deduce imediat din (8.6.34), (8.6.35) și din faptul că

$$\sum_{k=0}^{\infty} p(k) = p^+(1).$$

Observație. Este clar că teorema precedentă admite următoarea formulare echivalentă: Fie $(\tilde{x}_n)$ un șir de variabile aleatoare discrete avînd valori întregi nenegative, iar (F_n) și $(p^+_{\tilde{x}_n})$ respectiv șirul funcțiilor de repartiție și al funcțiilor generatoare corespunzătoare acestor variabile aleatoare. Dacă șirul $(p^+_{\tilde{x}_n}(z))$ converge către o funcție generatoare $p^+(z)$ pentru $|z| \leqslant 1$, atunci (F_n) converge în repartiție spre o funcție de repartiție F_0 discretă și, în plus, $p^+(z)$ este funcția generatoare corespunzătoare lui F_0.

Remarcăm totodată că dacă

$$\lim_n p^+_{\tilde{x}_n}(z) = p^+(z)$$

numai pentru $|z| < 1$, nu rezultă în mod necesar că $p^+(z)$ este o funcție generatoare corespunzătoare unei variabile aleatoare.

Scanned with OKEN Scanner

Într-adevăr, să considerăm șirul variabilelor aleatoare discrete $(\tilde{x}_n)$, unde $\tilde{x}_n$ are repartiția

$$\tilde{x}_n : \begin{pmatrix} 0 & n \\ 1/2 & 1/2 \end{pmatrix}.$$

Avem

$$p^{+}_{\tilde{x}_n}(z) = (1+z^n)/2$$

și deci

$$\lim_n p^{+}_{\tilde{x}_n}(z) = 1/2$$

pentru $|z| < 1$.

Dar

$$\sum_{k=0}^{\infty} p(k) = 1/2$$

și deci $p^{+}(z)$ nu este funcția generatoare a unei variabile aleatoare discrete (nu satisface condiția (8.6.28)).

Scanned with OKEN Scanner

Capitolul 9

LEGEA NUMERELOR MARI

9.1. INTRODUCERE

Observații empirice asupra fenomenelor aleatoare au condus la concluzia că, în studierea unui anumit fenomen particular, nu putem disocia esența sa de toate aspectele colaterale care fac dificilă depistarea legilor care pot fi puse în evidență de observarea unui număr mare de fenomene similare. Factorii colaterali, care apar numai într-o singură observație, se compensează reciproc cînd se consideră un număr mare de observații, putîndu-se pune în evidență esența fenomenelor considerate.

Aceste observații empirice pot fi pe deplin explicate pe baza teoriei probabilităților. Rezultatul teoretic care explică constatările empirice la care ne-am referit mai sus este cunoscut sub numele de *legea numerelor mari*. Sub acest nume vom înțelege de fapt o serie de propoziții care afirmă că, cu o probabilitate apropiată de 1, se va produce un anumit eveniment, eveniment care depinde de un număr arbitrar de mare de evenimente, fiecare dintre ele avînd o influență mică asupra fenomenului considerat.

Pentru a înțelege mai ușor esența problemelor expuse mai sus, să dăm un exemplu simplu. La aruncarea unei monede perfecte, probabilitatea de apariție a uneia din fețe este 0,5. Nu putem însă ști, înainte de a efectua aruncarea, care dintre cele două fețe va apărea. Să considerăm un număr n de aruncări ale acestei monede. În k dintre aceste aruncări apare stema, iar în celelalte $n-k$ apare valoarea. Raportul k/n reprezintă *frecvența (relativă)* de apariție a stemei (evident, raportul $(n-k)/n$ reprezintă frecvența de apariție a valorii). Dacă n este mare, este de așteptat ca frecvența k/n să fie „destul de apropiată" de 0,5. Statisticianul englez K. Pearson a aruncat o monedă de $n = 12\,000$ ori și a obținut stema de $k = 6\,019$ ori, ceea ce dă o frecvență $k/n = 0{,}5016$. Același statistician, într-un alt experiment, a aruncat o monedă de $n = 24\,000$ ori și a obținut

stema de $k = 12\,012$ ori, ceea ce dă o frecvență de $k/n = 0{,}5005$. În ambele cazuri, diferența dintre frecvență și probabilitate este destul de mică. Totuși, așa cum vom vedea în curînd, nu putem afirma că frecvența k/n „converge" la 0,5 cînd n crește indefinit, așa cum s-ar putea crede la prima vedere.

Exemplul de mai sus se poate generaliza în modul următor. Dacă $(E, \mathcal{K}, P)$ este un cîmp borelian de probabilitate, fie $A \in \mathcal{K}$ cu $P(A) = p \in (0,1)$ și fie $(\tilde{x}_n)$ șirul de variabile aleatoare independente definite de

$$\tilde{x}_n(\omega) = \begin{cases} 1 \text{ pentru } \omega \in A, \\ 0 \text{ pentru } \omega \in A^c. \end{cases}$$

Dacă se fac experimente independente în urma cărora se poate produce fie A fie A^c, atunci este clar că $\tilde{x}_n$ se poate interpreta ca numărul (1 sau 0) de apariții ale lui A în experimentul de rang n. Rezultă deci că numărul de apariții ale lui A în n experimente este $\tilde{k} = \tilde{x}_1 + \ldots + \tilde{x}_n$, astfel încît frecvența de apariție a lui A este

$$\tilde{k}/n = (\tilde{x}_1 + \ldots + \tilde{x}_n)/n. \tag{9.1.1}$$

După cum am spus și mai sus, este de așteptat ca frecvența $\tilde{k}/n$ dată de (9.1.1) să fie „destul de aproape" de probabilitatea p.

9.2. FORMULAREA PROBLEMEI

Să considerăm șirul de variabile aleatoare $(\tilde{x}_n)$ pe cîmpul borelian de probabilitate $(E, \mathcal{K}, P)$. Să considerăm funcțiile $g_n : R^n \to R$ măsurabile și simetrice în argumentele lor. Vom nota prin $\tilde{y}_n$ variabila aleatoare $g_n(\tilde{x}_1, \ldots, \tilde{x}_n)$, adică

$$\tilde{y}_n(\omega) = g_n(\tilde{x}_1(\omega), \ldots, \tilde{x}_n(\omega)).$$

Definiția 9.2.1. *Spunem că șirul de variabile aleatoare* $(\tilde{x}_n)$ *este supus legii slabe a numerelor mari sau este slab stabil dacă există un șir de constante reale* (a_n) *astfel încît șirul de variabile aleatoare* $(\tilde{y}_n - a_n)$ *să conveargă în probabilitate spre zero.*

Scanned with OKEN Scanner

Definiția 9.2.2. *Vom spune că șirul de variabile aleatoare* $(\tilde{x}_n)$ *este supus legii tari a numerelor mari sau este tare stabil dacă există un șir de constante reale* (a_n) *astfel încît șirul de variabile aleatoare* $(\tilde{y}_n - a_n)$ *să conveargă aproape sigur spre zero.*

În cele ce urmează ne vom restrînge la cazul în care funcțiile $g_n : R^n \to R$ sînt de forma

$$g_n(x_1, \ldots, x_n) = (x_1 + \ldots x_n)/n,$$

iar (a_n) este șirul de numere reale dat de relația

$$a_n = (M(\tilde{x}_1) + \ldots + M(\tilde{x}_n))/n.$$

Evident, se presupune implicit că variabilele aleatoare $\tilde{x}_1, \ldots \tilde{x}_n$ sînt integrabile și deci mediile lor există.

Exemplul dat în § 9.1 se încadrează în limitele impuse mai sus. Într-adevăr, pentru acest caz avem

$$\tilde{y}_n = (\tilde{x}_1 + \ldots + \tilde{x}_n)/n = \tilde{k}/n.$$

Deoarece

$$M(\tilde{x}_i) = 1 \cdot p + 0 \cdot (1 - p) = p,$$

rezultă că

$$a_n = \frac{p + \ldots + p}{n} = np/n = p.$$

Cu alte cuvinte, șirul de variabile aleatoare $(\tilde{y}_n - a_n)$ este în acest caz $(\tilde{k}/n - p)$. După cum vom vedea acest șir este atît slab cît și tare stabil (vezi teoremele 9.3.6 și 9.4.6).

9.3. LEGEA SLABĂ A NUMERELOR MARI

Vom da în cele ce urmează diverse variante ale legii slabe a numerelor mari.

Teorema 9.3.1 *(Markov). Fie* $(\tilde{x}_n)$ *un șir de variabile aleatoare. Dacă*

$$\lim_n (1/n^2)\, V\left(\sum_{i=1}^{n} \tilde{x}_i\right) = 0, \qquad (9.3.1)$$

atunci $(\tilde{x}_n)$ *este supus legii slabe a numerelor mari.*

Demonstrație. Utilizînd inegalitatea lui Cebîșev (vezi § 5.4) obținem

$$P(|(1/n)\sum_{i=1}^{n}\tilde{x}_i(\omega) - (1/n)\sum_{i=1}^{n}M(\tilde{x}_i)| \geqslant \varepsilon) \leqslant \frac{1}{\varepsilon^2 n^2}V\left(\sum_{i=1}^{n}\tilde{x}_i\right).$$

Trecînd la limită pentru $n \to +\infty$ în această inegalitate, obținem rezultatul căutat.

Teorema 9.3.2. *Fie* $(\tilde{x}_n)$ *un șir de variabile aleatoare independente două cîte două. Dacă*

$$\lim_n (1/n^2)\sum_{i=1}^{n}V(\tilde{x}_i) = 0, \tag{9.3.2}$$

atunci $(\tilde{x}_n)$ *urmează legea slabă a numerelor mari.*

Demonstrație. Rezultatul este o consecință imediată a teoremei precedente, dacă se observă că în cazul de față avem

$$V\left(\sum_{i=1}^{n}\tilde{x}_i\right) = \sum_{i=1}^{n}V(\tilde{x}_i).$$

Teorema 9.3.3 *(Cebîșev). Dacă* $(\tilde{x}_n)$ *este un șir de variabile aleatoare independente două cîte două care au dispersii finite mărginite de aceeași constantă* C, *adică* $V(\tilde{x}_n) \leqslant C$ *pentru orice* n, *atunci* $(\tilde{x}_n)$ *urmează legea slabă a numerelor mari.*

Demonstrație. Rezultatul este o consecință imediată a teoremei precedente dacă ținem seama că

$$(1/n^2)\sum_{i=1}^{n}V(\tilde{x}_i) \leqslant (1/n^2)\,nC = C/n$$

și deci (9.3.2) este îndeplinită.

Teorema 9.3.4. *Dacă* $(\tilde{x}_n)$ *este un șir de variabile aleatoare independente două cîte două și identic repartizate (cu aceeași funcție de repartiție), dacă există* $M(\tilde{x}_n) = a$ *și dispersia* $V(\tilde{x}_n) \leqslant C$, *atunci șirul* $(\tilde{x}_n)$ *urmează legea slabă a numerelor mari.*

Demonstrație. Teorema este un caz particular al teoremei precedente.

Scanned with OKEN Scanner

Observație. Din această teoremă rezultă că, în ipotezele specificate, avem

$$\lim_n P(|(1/n)\sum_{i=1}^{n}\tilde{x}_i - a| \geqslant \varepsilon) = 0.$$

Această formulă dă o justificare a regulei mediei aritmetice, aplicată curent în măsurători. Dacă a este o mărime ce trebuie măsurată și se fac n măsurători care produc rezultatele $\tilde{x}_1, \ldots, \tilde{x}_n$, atunci drept valoare aproximativă a lui a se ia media aritmetică

$$(\tilde{x}_1 + \ldots \tilde{x}_n)/n \tag{9.3.3}$$

a rezultatelor obținute. Dacă $M(\tilde{x}_1) = \ldots = M(\tilde{x}_n) = a$ (adică dacă lipsesc erorile sistematice de măsurare), atunci legea numerelor mari ne asigură că, cu o probabilitate oricît de apropiată de 1, putem obține, pentru n suficient de mare, o medie aritmetică (9.3.3) oricît de apropiată de a.

Teorema 9.3.5 *(Poisson). Fie* $\tilde{\mathrm{x}}_{(\mathrm{n})}$ *numărul de apariții al unui eveniment* A *în* n *experimente independente, probabilitatea producerii lui* A *în experimentul de rang* k *fiind notată prin* $\mathrm{p_k}$. *În aceste condiții, pentru orice* $\varepsilon > 0$, *avem*

$$\lim_{\mathrm{n}} \mathrm{P}(|\tilde{\mathrm{x}}_{(\mathrm{n})}/\mathrm{n} - (\mathrm{p_1} + \ldots + \mathrm{p_n})/\mathrm{n}| \geqslant \varepsilon) = 0. \tag{9.3.4}$$

Demonstrație. Fie $\tilde{x}_k$ o variabilă aleatoare care ia valoarea 1 sau 0 după cum A apare sau nu în experimentul de rang k. Evident, variabilele aleatoare $(\tilde{x}_k)$ astfel definite sînt independente, iar

$$M(\tilde{x}_k) = 1 \cdot p_k + 0 \cdot (1 - p_k) = p_k$$

și

$$V(\tilde{x}_k) = p_k - p_k^2 = p_k(1 - p_k) \leqslant 1/4.$$

Rezultă imediat că avem

$$\tilde{x}_{(n)} = \tilde{x}_1 + \ldots + \tilde{x}_n \tag{9.3.5}$$

și deci

$$M(\tilde{x}_{(n)}) = M(\tilde{x}_1) + \ldots + M(\tilde{x}_n) = p_1 + \ldots + p_n. \tag{9.3.6}$$

Scanned with OKEN Scanner

Observație. Din această teoremă rezultă că, în ipotezele specificate, avem

$$\lim_n P(|(1/n)\sum_{i=1}^{n}\tilde{x}_i - a| \geqslant \varepsilon) = 0.$$

Această formulă dă o justificare a regulei mediei aritmetice, aplicată curent în măsurători. Dacă a este o mărime ce trebuie măsurată și se fac n măsurători care produc rezultatele $\tilde{x}_1, \ldots, \tilde{x}_n$, atunci drept valoare aproximativă a lui a se ia media aritmetică

$$(\tilde{x}_1 + \ldots \tilde{x}_n)/n \tag{9.3.3}$$

a rezultatelor obținute. Dacă $M(\tilde{x}_1) = \ldots = M(\tilde{x}_n) = a$ (adică dacă lipsesc erorile sistematice de măsurare), atunci legea numerelor mari ne asigură că, cu o probabilitate oricît de apropiată de 1, putem obține, pentru n suficient de mare, o medie aritmetică (9.3.3) oricît de apropiată de a.

Teorema 9.3.5 *(Poisson). Fie* $\tilde{\mathrm{x}}_{(\mathrm{n})}$ *numărul de apariții al unui eveniment* A *în* n *experimente independente, probabilitatea producerii lui* A *în experimentul de rang* k *fiind notată prin* $\mathrm{p_k}$. *În aceste condiții, pentru orice* $\varepsilon > 0$, *avem*

$$\lim_{\mathrm{n}} \mathrm{P}(|\tilde{\mathrm{x}}_{(\mathrm{n})}/\mathrm{n} - (\mathrm{p_1} + \ldots + \mathrm{p_n})/\mathrm{n}| \geqslant \varepsilon) = 0. \tag{9.3.4}$$

Demonstrație. Fie $\tilde{x}_k$ o variabilă aleatoare care ia valoarea 1 sau 0 după cum A apare sau nu în experimentul de rang k. Evident, variabilele aleatoare $(\tilde{x}_k)$ astfel definite sînt independente, iar

$$M(\tilde{x}_k) = 1 \cdot p_k + 0 \cdot (1 - p_k) = p_k$$

și

$$V(\tilde{x}_k) = p_k - p_k^2 = p_k(1 - p_k) \leqslant 1/4.$$

Rezultă imediat că avem

$$\tilde{x}_{(n)} = \tilde{x}_1 + \ldots + \tilde{x}_n \tag{9.3.5}$$

și deci

$$M(\tilde{x}_{(n)}) = M(\tilde{x}_1) + \ldots + M(\tilde{x}_n) = p_1 + \ldots + p_n. \tag{9.3.6}$$

Scanned with OKEN Scanner

Evident, șirul $(\tilde{x}_n)$ de variabile aleatoare verifică ipotezele teoremei 9.3.3 și deci $(\tilde{x}_n)$ urmează legea slabă a numerelor mari, adică, pentru orice $\varepsilon > 0$, avem

$$\lim_n P(|(1/n)\sum_{i=1}^{n} \tilde{x}_i - (1/n)\sum_{i=1}^{n} M(\tilde{x}_i)| \geqslant \varepsilon) = 0. \qquad (9.3.7)$$

Este clar că (9.3.4) rezultă imediat din (9.3.5), (9.3.6) și (9.3.7).

Teorema 9.3.6 *(Bernoulli). Fie* $\tilde{x}_{(n)}$ *numărul de apariții ale unui eveniment* A *în* n *experimente independente și* p *probabilitatea producerii lui* A *în fiecare experiment. În aceste condiții,*

$$\lim_n P(|\tilde{x}_{(n)}/n - p| \geqslant \varepsilon) = 0$$

pentru orice $\varepsilon > 0$.

Demonstrație. Rezultatul se obține ca un caz particular din teorema precedentă ($p_n = p$ pentru orice n).

Observație. Conform cu cele spuse în § 9.1, raportul $\tilde{x}_{(n)}/n$ reprezintă frecvența (relativă) de apariție a evenimentului A în n experimente independente. Teorema lui Bernoulli afirmă deci că frecvența converge în probabilitate către probabilitatea evenimentului respectiv.

Teorema 9.3.7 *(Hincin). Dacă* $(\tilde{x}_n)$ *este un șir de variabile aleatoare independente și identic repartizate cu media finită* $(M(\tilde{x}_n) = a < +\infty)$, *atunci* $(\tilde{x}_n)$ *este supus legii slabe a numerelor mari.*

Demonstrație. Fie F funcția de repartiție a variabilei aleatoare $\tilde{x}_n$, $n \geqslant 1$, și fie $\delta > 0$ fixat. Deoarece $M(\tilde{x}_n) = a$ este finit, rezultă că putem scrie

$$a = \int_{-\infty}^{\infty} x \, dF(x) = \lim_{n\to\infty} \int_{-\delta n}^{\delta n} x \, dF(x)$$

și deci, pentru orice $\varepsilon > 0$, există $N(\varepsilon)$ cu proprietatea că pentru orice $n \geqslant N(\varepsilon)$ avem

$$|a_n - a| < \varepsilon, \qquad (9.3.8)$$

unde

$$a_n = \int_{-\delta n}^{\delta n} x \, dF(x). \tag{9.3.9}$$

Să fixăm acum n și să considerăm variabilele aleatoare $\tilde{y}_k$ și $\tilde{z}_k$, definite pentru $k = 1, 2, \ldots, n$ prin

$$\tilde{y}_k = \begin{cases} \tilde{x}_k & \text{dacă } |\tilde{x}_k| < \delta n, \\ 0 & \text{dacă } |\tilde{x}_k| \geqslant \delta n \end{cases} \tag{9.3.10}$$

și

$$\tilde{z}_k = \begin{cases} 0 & \text{dacă } |\tilde{x}_k| < \delta n, \\ \tilde{x}_k & \text{dacă } |\tilde{x}_k| \geqslant \delta n. \end{cases} \tag{9.3.11}$$

Evident, pentru fiecare k, $1 \leqslant k \leqslant n$, avem

$$\tilde{x}_k = \tilde{y}_k + \tilde{z}_k \tag{9.3.12}$$

Este de asemenea clar că media lui $\tilde{y}_k$ există și este a_n. Mai mult, există și dispersia lui $\tilde{y}_k$, deoarece

$$V(\tilde{y}_k) = \int_{-\delta n}^{\delta n} x^2 \, dF(x) - a_n^2 \leqslant \delta n \int_{-\delta n}^{\delta n} |x| \, dF(x) \leqslant \delta n A, \tag{9.3.13}$$

unde

$$A = \int_{-\infty}^{\infty} |x| \, dF(x). \tag{9.3.14}$$

Din (9.3.13), aplicînd inegalitatea lui Cebîșev, obținem

$$P\left(\left| (1/n) \sum_{k=1}^{n} \tilde{y}_k - a_n \right| \geqslant \varepsilon \right) \leqslant \delta A / \varepsilon^2. \tag{9.3.15}$$

În sfîrșit, din (9.3.8) și (9.3.15) rezultă

$$P\left(\left| (1/n) \sum_{k=1}^{n} \tilde{y}_k - a \right| \geqslant 2\varepsilon \right) \leqslant \delta A / \varepsilon^2. \tag{9.3.16}$$

Scanned with OKEN Scanner

Pe de altă parte, să observăm că

$$P(\tilde{z}_k \neq 0) = P(|\tilde{x}_k| \geqslant \delta n) = \int\limits_{|x| \geqslant \delta n} \mathrm{d}F(x) \leqslant$$
$$\leqslant (1/\delta n) \int\limits_{|x| \geqslant \delta n} |x| \quad \mathrm{d}F(x). \tag{9.3.17}$$

Deoarece valoarea medie există și este finită, rezultă că pentru n suficient de mare putem scrie

$$\int\limits_{|x| \geqslant \delta n} |x| \quad \mathrm{d}F(x) < \delta^2$$

și deci (9.3.17) devine

$$P(\tilde{z}_k \neq 0) \leqslant \delta/n. \tag{9.3.18}$$

Prin urmare avem

$$P\left(\sum_{k=1}^{n} \tilde{z}_k \neq 0\right) \leqslant \sum_{k=1}^{n} P(\tilde{z}_k \neq 0) \leqslant n \cdot (\delta/n) = \delta. \tag{9.3.19}$$

Din (9.3.16) și (9.3.19) obținem atunci

$$P\left(\left|(1/n)\sum_{k=1}^{n} \tilde{x}_k - a\right| \geqslant 2\varepsilon\right) \leqslant P\left(\left|(1/n)\sum_{k=1}^{n} \tilde{y}_k - a\right| \geqslant 2\varepsilon\right) +$$
$$+ P\left(\sum_{k=1}^{n} \tilde{z}_k \neq 0\right) \leqslant \delta A/\varepsilon^2 + \delta. \tag{9.3.20}$$

Deoarece δ și ε sînt arbitrari, membrul drept din (9.3.20) poate fi făcut oricît de mic și rezultatul se deduce imediat.

Vom da în încheiere un criteriu de verificare a stabilității slabe.

Scanned with OKEN Scanner

Teorema 9.3.8. *Fie* $(\tilde{x}_n)$ *un șir de variabile aleatoare. Condiția necesară și suficientă ca* $(\tilde{x}_n)$ *să fie supus legii slabe a numerelor mari este ca*

$$\lim_n M\left[\frac{(\tilde{y}_1 + \ldots + \tilde{y}_n)^2}{n^2 + (\tilde{y}_1 + \ldots + \tilde{y}_n)^2}\right] = 0, \tag{9.3.21}$$

unde

$$\tilde{y}_k = \tilde{x}_k - M(\tilde{x}_k),\ 1 \leqslant k \leqslant n. \tag{9.3.22}$$

Demonstrație. Să considerăm variabilele aleatoare $\tilde{z}_n$ definite de relația

$$\tilde{z}_n = (\tilde{y}_1 + \ldots + \tilde{y}_n)/n$$

și să notăm prin $G_n(x)$ funcția de repartiție a variabilei aleatoare $\tilde{z}_n$.

Necesitatea condiției rezultă în modul următor. Pentru orice $\varepsilon > 0$ avem

$$P(|\tilde{z}_n| \geqslant \varepsilon) = \int_{|x| \geqslant \varepsilon} dG_n(x) \geqslant \int_{|x| \geqslant \varepsilon} \frac{x^2}{1+x^2}\, dG_n(x) =$$

$$= \int_{-\infty}^{\infty} \frac{x^2}{1+x^2}\, dG_n(x) - \int_{|x| < \varepsilon} \frac{x^2}{1+x^2}\, dG_n(x) \geqslant$$

$$\geqslant \int_{-\infty}^{\infty} \frac{x^2}{1+x^2}\, dG_n(x) - \varepsilon^2 = M\left(\frac{\tilde{z}_n^2}{1+\tilde{z}_n^2}\right) - \varepsilon^2$$

și deci

$$0 \leqslant M\left(\frac{\tilde{z}_n^2}{1+\tilde{z}_n^2}\right) \leqslant \varepsilon^2 + P(|\tilde{z}_n| \geqslant \varepsilon),$$

de unde se obține imediat condiția (9.3.21).

Scanned with OKEN Scanner

Suficiența condiției rezultă din șirul de inegalități:

$$P(|\tilde{z}_n| \geqslant \varepsilon) = \int\limits_{|x| \geqslant \varepsilon} dG_n(x) \leqslant \frac{1+\varepsilon^2}{\varepsilon^2} \int\limits_{|x| \geqslant \varepsilon} \frac{x^2}{1+x^2} dG_n(x) \leqslant$$

$$\leqslant \frac{1+\varepsilon^2}{\varepsilon^2} \int\limits_{-\infty}^{\infty} \frac{x^2}{1+x^2} dG_n(x) = \frac{1+\varepsilon^2}{\varepsilon^2} M\left(\frac{\tilde{z}_n^2}{1+\tilde{z}_n^2}\right)$$

9.4. LEGEA TARE A NUMERELOR MARI

Vom da mai întîi o inegalitate datorată lui Kolmogorov, care va constitui instrumentul principal în obținerea de condiții suficiente ca un șir de variabile aleatoare să fie tare stabil.

Teorema 9.4.1 *(Inegalitatea lui Kolmogorov). Fie* $(\tilde{x}_n)$ *un șir de variabile aleatoare independente. Dacă* $M(\tilde{x}_j) = m_j$ *și* $V(\tilde{x}_j) = \sigma_j^2$ *există pentru fiecare* j, $1 \leqslant j \leqslant n$, *atunci*

$$P\left(\max_{1 \leqslant j \leqslant n} \left| \sum_{j=1}^{n} (\tilde{x}_j - m_j) \right| \geqslant \varepsilon\right) \leqslant \varepsilon^{-2} \sum_{j=1}^{n} \sigma_j^2 \qquad (9.4.1)$$

pentru orice $\varepsilon > 0$.

Demonstrație. Să considerăm variabilele aleatoare

$$\tilde{y}_j = \tilde{x}_j - m_j, \qquad 1 \leqslant j \leqslant n$$

și

$$\tilde{z}_j = \tilde{y}_1 + \ldots + \tilde{y}_j, \qquad 1 \leqslant j \leqslant n.$$

Cu aceste notații, inegalitatea (9.4.1) se poate scrie sub forma

$$P\left(\max_{1 \leqslant j \leqslant n} |\tilde{z}_j(\omega)| \geqslant \varepsilon\right) \leqslant \varepsilon^{-2} \sum_{j=1}^{n} \sigma_j^2. \qquad (9.4.2)$$

Scanned with OKEN Scanner

Fie evenimentele $A_1, A_2, \ldots, A_{n+1}$ definite astfel:

$$A_1 = \{\omega : |\tilde{z}_1(\omega)| \geqslant \varepsilon\};$$

$$A_j = \{\omega : |\tilde{z}_i(\omega)| < \varepsilon \text{ pentru } 1 \leqslant i < j-1 \text{ și } |\tilde{z}_j(\omega)| \geqslant \varepsilon\},\ 2 \leqslant j \leqslant n;$$

$$A_{n+1} = \{\omega : |\tilde{z}_j(\omega)| < \varepsilon \text{ pentru } 1 \leqslant j \leqslant n\}.$$

Este clar că $A_1, \ldots, A_n$ sînt incompatibile două cîte două și

$$\left\{\omega : \max_{1 \leqslant j \leqslant n} |\tilde{z}_j(\omega)| \geqslant \varepsilon\right\} = \bigcup_{j=1}^{n} A_j.$$

Rezultă atunci că

$$P\left(\max_{1 \leqslant j \leqslant n} |\tilde{z}_j| \geqslant \varepsilon\right) = P\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} P(A_j). \qquad (9.4.3)$$

Să observăm de asemenea că $A_1, A_2, \ldots, A_n, A_{n+1}$ formează o desfacere finită și deci putem scrie

$$V(\tilde{z}_n) = M(\tilde{z}_n^2) = \sum_{j=1}^{n+1} P(A_j)\, M(\tilde{z}_j^2 \mid A_j) \geqslant \geqslant \sum_{j=1}^{n} P(A_j)\, M(\tilde{z}_j^2) \mid A_j). \qquad (9.4.4)$$

Pentru fiecare j, $1 \leqslant j \leqslant n$, avem

$$M\left(\tilde{z}_n^2 \,\middle|\, A_j\right) = M\left(\left(\tilde{z}_j + \sum_{i=j+1}^{n} \tilde{y}_i\right)^2 \,\middle|\, A_j\right) =$$

$$= M\left(\tilde{z}_j^2 + \sum_{i=j+1}^{n} \tilde{y}_i^2 + 2\sum_{i=j+1}^{n} \tilde{z}_j \tilde{y}_i + 2 \sum_{\substack{i,k=j+1 \\ i>k}}^{n} \tilde{y}_i \tilde{y}_k \,\middle|\, A_j\right) \geqslant$$

$$\geqslant M\left(\tilde{z}_j^2 + 2\sum_{i=j+1}^{n} \tilde{z}_j \tilde{y}_i + 2 \sum_{\substack{i,k=j+1 \\ i>k}}^{n} \tilde{y}_i \tilde{y}_k \,\middle|\, A_j\right). \qquad (9.4.5)$$

Este clar că variabilele aleatoare $\tilde{y}_i$, $i \geqslant j+1$, sînt independente între ele și independente de $\tilde{z}_j$, dacă P este înlocuit prin $P(\cdot \mid A_j)$. Rezultă atunci că

$$M(\tilde{z}_j \tilde{y}_i \mid A_j) = M(\tilde{z}_j \mid A_j)\, M(\tilde{y}_i \mid A_j) = 0, \quad j+1 \leqslant i \leqslant n, \tag{9.4.6}$$

și

$$M(\tilde{y}_i \tilde{y}_k \mid A_j) = M(\tilde{y}_i \mid A_j)\, M(\tilde{y}_k \mid A_j) = 0, \quad i > k \geqslant j+1. \tag{9.4.7}$$

În sfîrșit, să observăm că

$$M(\tilde{z}_j^2 \mid A_j) \geqslant \varepsilon^2, \quad 1 \leqslant j \leqslant n, \tag{9.4.8}$$

dacă ținem seama că pentru $\omega \in A_j$ avem $\tilde{z}_j^2(\omega) \geqslant \varepsilon$.

Utilizînd acum (9.4.5), (9.4.6), (9.4.7) și (9.4.8) obținem

$$M(\tilde{z}_n^2 \mid A_j) \geqslant \varepsilon^2, \quad 1 \leqslant j \leqslant n. \tag{9.4.9}$$

În sfîrșit, din (9.4.3), (9.4.4) și (9.4.9) obținem

$$V(\tilde{z}_n) \geqslant \varepsilon^2 \sum_{j=1}^{n} P(A_j) = \varepsilon^2 P\left(\max_{1 \leqslant j \leqslant n} |\tilde{z}_j| \geqslant \varepsilon\right),$$

de unde rezultă (9.4.2), dacă ținem seama și de faptul că

$$V(\tilde{z}_n) = V\left(\sum_{j=1}^{n} \tilde{y}_j\right) = \sum_{j=1}^{n} V(\tilde{y}_j) = \sum_{j=1}^{n} V(\tilde{x}_j) = \sum_{j=1}^{n} \sigma_j^2.$$

Observație. Este clar că inegalitatea (9.4.1) se poate scrie sub forma echivalentă

$$P\left(\left|\sum_{i=1}^{j} (\tilde{x}_i - m_i)\right| \leqslant \varepsilon \text{ pentru } j = 1, 2, \ldots, n\right) \geqslant \tag{9.4.10}$$

$$\geqslant 1 - (1/\varepsilon^2) \sum_{j=1}^{n} \sigma_j^2.$$

Într-adevăr, avem evident

$$\left\{\omega : \max_{1\leqslant j\leqslant n}\left|\sum_{j=1}^{n}(\widetilde{x}_j(\omega)-m_j)\right|\geqslant\varepsilon\right\}=$$

$$=\left\{\omega:\left|\sum_{i=1}^{j}(\widetilde{x}_i(\omega)-m_i)\right|\leqslant\varepsilon,\ 1\leqslant j\leqslant n\right\}^c.$$

Teorema care urmează dă o condiție suficientă de convergență aproape sigură, condiție ce va fi utilizată, împreună cu inegalitatea lui Kolmogorov, la obținerea de condiții suficiente de stabilitate tare a șirurilor de variabile aleatoare.

Teorema 9.4.2. *Fie* $(\widetilde{x}_n)_{n\geqslant 0}$ *un șir de variabile aleatoare. Dacă pentru fiecare întreg pozitiv* k,

$$\sum_{n=1}^{\infty}P(|\widetilde{x}_n-\widetilde{x}_0|\geqslant 1/k)<+\infty, \tag{9.4.11}$$

atunci $(\widetilde{x}_n)$ *converge aproape sigur spre* $\widetilde{x}_0$.

Demonstrație. Fie A_n^k evenimentul definit de

$$A_n^k=\{\omega:|\widetilde{x}_n(\omega)-\widetilde{x}_0(\omega)|\geqslant 1/k\}$$

și fie

$$B_n^k=\bigcup_{j=0}^{\infty}A_{n+j}^k.$$

Deoarece

$$P(B_n^k)\leqslant\sum_{j=0}^{\infty}P(A_{n+j}^k)=\sum_{i=n}^{\infty}P(A_i^k)=\sum_{i=n}^{\infty}P(|\widetilde{x}_i-\widetilde{x}_0|\geqslant 1/k),$$

din (9.4.11) rezultă

$$\lim_n P(B_n^k)=0$$

pentru orice k.

Este clar acum că

$$P\left(\bigcup_{k=1}^{\infty}\bigcap_{n=1}^{\infty}B_n^k\right)\leqslant\sum_{k=1}^{\infty}P\left(\bigcap_{n=1}^{\infty}B_n^k\right)\leqslant\sum_{k=1}^{\infty}P(B_n^k).$$

Scanned with OKEN Scanner

Din faptul că

$$P\left(\bigcap_{n=1}^{\infty} B_n^k\right) \leqslant P(B_n^k)$$

rezultă prin trecere la limită că

$$P\left(\bigcap_{n=1}^{\infty} B_n^k\right) = 0.$$

Urmează atunci imediat că

$$P\left(\bigcup_{k=1}^{\infty} \bigcap_{n=1}^{\infty} B_n^k\right) \leqslant \sum_{k=1}^{\infty} P\left(\bigcap_{n=1}^{\infty} B_n^k\right) = 0,$$

adică

$$P\left(\bigcup_{k=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{j=0}^{\infty} \{\omega : |\tilde{x}_{n+j}(\omega) - \tilde{x}_0(\omega)| \geqslant 1/k\}\right) = 0,$$

de unde se obține imediat că $(\tilde{x}_n)$ converge aproape sigur spre $\tilde{x}_0$.

Observație. Se poate arăta ușor, reluînd demonstrația de mai sus, că are loc un rezultat ceva mai tare decît cel enunțat în teorema precedentă: dacă există un șir de numere întregi $1 = n_1 < n_2 < \ldots$ astfel încît seria

$$\sum_{j=1}^{\infty} P\left(\max_{n_j \leqslant n < n_{j+1}} |\tilde{x}_n - \tilde{x}_0| \geqslant 1/k\right) \qquad (9.4.12)$$

să conveargă pentru orice întreg pozitiv k, atunci $(\tilde{x}_n)$ converge aproape sigur spre $\tilde{x}_0$.

Teorema 9.4.3. *(Kolmogorov). Dacă șirul de variabile aleatoare independente* $(\tilde{x}_n)$ *satisface condiția*

$$\sum_{n=1}^{\infty} (1/n^2)\, V(\tilde{x}_n) < +\infty, \qquad (9.4.13)$$

atunci $(\tilde{x}_n)$ *este supus legii tari a numerelor mari.*

Demonstrație. Să considerăm variabilele aleatoare

$$\tilde{y}_n = \sum_{j=1}^{n} (\tilde{x}_j - M(\tilde{x}_j))$$

și

$$\tilde{z}_n = (1/n)\, \tilde{y}_n.$$

Pentru a demonstra teorema este suficient să arătăm că $(\tilde{z}_n)$ converge aproape sigur spre zero. Să considerăm pentru aceasta $\varepsilon > 0$, un întreg pozitiv m și să definim probabilitățile $P(m, \varepsilon)$ prin

$$P(m, \varepsilon) = P\left(\max_n |\tilde{z}_n| \leqslant \varepsilon,\ 2^m \leqslant n < 2^{m+1}\right).$$

Deoarece putem scrie

$$P(m, \varepsilon) \leqslant P\left(\max_n |\tilde{y}_n| \leqslant 2^m \varepsilon,\ 2^m \leqslant n < 2^{m+1}\right),$$

din inegalitatea lui Kolmogorov obținem

$$P(m, \varepsilon) \leqslant (2^m \varepsilon)^{-2} \sum_{j=1}^{2^{m+1}-1} V(\tilde{x}_j)$$

și deci

$$\sum_{m=1}^{\infty} P(m, \varepsilon) \leqslant \varepsilon^{-2} \sum_{m=1}^{\infty} 2^{-2m} \sum_{j=1}^{2^{2m+1}-1} V(\tilde{x}_j) = \varepsilon^{-2} \sum_{j=1}^{\infty} V(\tilde{x}_j) \sum_{m \in M_j} 2^{-2m},$$

unde

$$M_j = \{m : 2^{m+1} > j\}.$$

Dacă se determină, pentru fiecare j, numărul m_j cu proprietatea

$$2^{m_j} \leqslant j < 2^{1+m_j},$$

atunci avem evident

$$\sum_{m \in M_j} 2^{-2m} = \sum_{m=m_j}^{\infty} 2^{-2m} = \frac{4}{3}\, 2^{-2m_j} < \frac{16}{3j^2},$$

ceea ce arată că

$$\sum_{m=1}^{\infty} P(m, \varepsilon) \leqslant \frac{16}{3\varepsilon^2} \sum_{j=1}^{\infty} \frac{V(\tilde{x}_j)}{j^2}.$$

Scanned with OKEN Scanner

Ținînd seama de condiția (9.4.13), ultima inegalitate ne conduce la concluzia că seria $\sum_{m=1}^{\infty} P(m, \varepsilon)$ este convergentă și deci, conform teoremei 9.4.2 (vezi și forma (9.4.12) a rezultatului), rezultă că $(\widetilde{z}_n)$ converge aproape sigur spre zero și astfel teorema este demonstrată.

Teorema 9.4.4. *Fie* $(\widetilde{x}_n)$ *un șir de variabile aleatoare independente. Dacă dispersiile variabilelor aleatoare* $(\widetilde{x}_n)$ *sînt mărginite de aceeași constantă reală, atunci* $(\widetilde{x}_n)$ *urmează legea tare a numerelor mari.*

Demonstrație. Fie $C > 0$ cu proprietatea că $V(\widetilde{x}_j) \leqslant C$ pentru orice j. Atunci

$$\sum_{n=1}^{\infty} \frac{V(\widetilde{x}_n)}{n^2} \leqslant \sum_{n=1}^{\infty} \frac{C}{n^2} < +\infty$$

și rezultatul se deduce din teorema precedentă.

Observație. Această teoremă este analoagă teoremei lui Cebîșev (vezi teorema 9.3.3). Așa cum teorema lui Poisson (vezi teorema 9.3.6) se deduce imediat din teorema lui Cebîșev, putem deduce analoage ale teoremelor lui Poisson și Bernoulli pentru legea tare a numerelor mari. Mai exact, putem obține rezultatele următoare.

Teorema 9.4.5. *Fie* $\widetilde{x}_{(n)}$ *numărul de apariții ale unui eveniment* A *în* n *experimente independente, probabilitatea producerii lui* A *în experimentul de rang* k *fiind notată prin* p_k. *În aceste condiții avem, cu probabilitatea* 1,

$$\lim_n (\widetilde{x}_{(n)}/n - (p_1 + \ldots + p_n)/n) = 0.$$

Teorema 9.4.6 *(Borel). Fie* $\widetilde{x}_{(n)}$ *numărul de apariții ale unui eveniment* A *în* n *experimente independente, probabilitatea producerii lui* A *într-un singur experiment fiind* p.

În aceste condiții, cu probabilitatea 1, *avem*

$$\lim_n (\widetilde{x}_{(n)}/n) = p. \qquad (9.4.14)$$

În sfîrșit, să observăm că teorema lui Hincin (vezi teorema 9.3.7) dă o condiție suficientă pentru ca un șir de variabile aleatoare independente și identic repartizate să urmeze legea slabă a numerelor mari. Vom prezenta în continuare o teoremă a lui Kolmogorov, care arată că această condiție este atît suficientă cît și necesară în cazul legii tari a numerelor mari.

Teorema 9.4.7 *(Kolmogorov). Dacă* $(\tilde{x}_n)$ *este un șir de variabile aleatoare independente și identic repartizate, atunci condiția necesară și suficientă ca* $(\tilde{x}_n)$ *să urmeze legea tare a numerelor mari este ca mediile* $M(\tilde{x}_n) = a$ *să existe și să fie finite.*

Demonstrație. Fie $F(x)$ funcția de repartiție a variabilelor aleatoare $\tilde{x}_n$, $n \geqslant 1$. Deoarece media există, rezultă că integrala

$$A = \int_{-\infty}^{\infty} |x| \, dF(x) \tag{9.4.15}$$

există și este finită.

Să considerăm variabilele aleatoare $\tilde{y}_n$ definite în modul următor:

$$\tilde{y}_n = \begin{cases} \tilde{x}_n & \text{pentru } |\tilde{x}_n| \leqslant n, \\ 0 & \text{pentru } |\tilde{x}_n| > n. \end{cases}$$

Este clar că

$$V(\tilde{y}_n) \leqslant M(\tilde{y}_n^2) = \int_{-n}^{n} x^2 \, dF(x) \leqslant \sum_{k=0}^{n} (k+1)^2 P(k < |\tilde{x}_j| \leqslant k+1)$$

și deci

$$\sum_{n=1}^{\infty} V(\tilde{y}_n)/n^2 \leqslant \sum_{n=1}^{\infty} \sum_{k=0}^{n} ((k+1)^2/n^2) \, P(k < |\tilde{x}_j| \leqslant k+1) =$$

$$= \sum_{k=1}^{\infty} (k+1)^2 P(k < |\tilde{x}_j| \leqslant k+1) \sum_{n=k}^{\infty} n^{-2} +$$

$$+ P(0 < |\tilde{x}_j| < 1) \sum_{n=1}^{\infty} n^{-2}.$$

Deoarece

$$\sum_{n=k}^{\infty} n^{-2} < k^{-2} + k^{-1} < 2/k,$$

rezultă

$$\sum_{n=1}^{\infty} V(\tilde{y}_n)/n^2 \leqslant \sum_{k=1}^{\infty} (2(k+1)^2/k)\, P(k < |\tilde{x}_j| \leqslant k+1) +$$

$$+ 2P(0 < |x_j| < 1) \leqslant 2 \sum_{k=1}^{\infty} kP\,(k < |\tilde{x}_j| \leqslant k+1) + 8 \leqslant$$

$$\leqslant 8 + 2 \sum_{k=0}^{\infty} \int_{k < |x| \leqslant k+1} x\, dF(x) \leqslant 8 + 2A,$$

ceea ce arată că $(\tilde{y}_n)$ urmează legea tare a numerelor mari, conform teoremei 9.4.3.

Pentru a termina demonstrația trebuie să mai arătăm că șirul $(\tilde{x}_n)$ urmează de asemenea legea tare a numerelor mari. Pentru aceasta este suficient să arătăm că probabilitatea ca $\tilde{x}_n$ să fie diferit de $\tilde{y}_n$ pentru $n \geqslant m$ tinde spre zero cînd $m \to +\infty$. Într-adevăr,

$$P(\tilde{x}_n \neq \tilde{y}_n \text{ pentru } n \geqslant m) \leqslant \sum_{n=m}^{\infty} P(\tilde{x}_n \neq \tilde{y}_n) = \sum_{n=m}^{\infty} P(|\tilde{x}_n| > n) =$$

$$= \sum_{n=m}^{\infty} \sum_{k=n}^{\infty} P(k < |\tilde{x}_n| \leqslant k+1) =$$

$$= \sum_{n=m}^{\infty} (n - m + 1)\, P(n \leqslant |\tilde{x}_n| < n+1) \leqslant$$

$$\leqslant \sum_{n=m}^{\infty} n \int_{n \leqslant |x| < n+1} dF(x) \leqslant \sum_{n=m}^{\infty} \int_{n \leqslant |x| < n+1} |x|\, dF(x) =$$

$$= \int_{|x| \geqslant m} |x|\, dF(x). \qquad (9.4.16)$$

Deoarece integrala (9.4.15) există și este finită, rezultă că pentru m suficient de mare integrala

$$\int_{|x|\geq m} |x| \, \mathrm{d}F(x)$$

poate fi făcută oricît de mică. Din (9.4.16) rezultă atunci ceea ce vroiam să arătăm și astfel teorema este complet demonstrată; reciproca se verifică în mod evident.

Scanned with OKEN Scanner

Capitolul 10

REPARTIȚIA NORMALĂ ȘI TEOREMA LIMITĂ CENTRALĂ

10.1. INTRODUCERE

Să considerăm o variabilă aleatoare $\tilde{x}$ pe cîmpul borelian de probabilitate $(E, \mathfrak{K}, P)$. Am văzut că „comportarea" variabilei aleatoare $\tilde{x}$ este descrisă de *repartiția** sa $P \circ \tilde{x}^{-1}$.

Uneori, această comportare este dată într-un alt mod. De exemplu, putem da repartiția lui x prin funcția de repartiție F, sau prin funcția de repartiție reciprocă G sau încă prin funcția H definită pentru orice interval $[x, y) \subset R$ prin relația

$$H([x, y)) = P(x \leqslant \tilde{x} < y).$$

De asemenea, putem da repartiția prin funcția caracteristică $\varphi_{\tilde{x}}$.

O variabilă aleatoare discretă $\tilde{x}$ poate fi dată prin funcția sa de masă, iar o variabilă aleatoare continuă poate fi dată prin densitatea sa de probabilitate.

Evident, aceleași observații sînt valabile pentru vectorii aleatori.

Studiul repartițiilor care au fost întîlnite în studiul fenomenelor aleatoare apărute în diverse probleme practice au dus la punerea în evidență a unor repartiții care se întîlnesc mai frecvent și pe care le vom numi *repartiții clasice*. Dintre aceste repartiții, cea mai importantă este *repartiția normală*, pusă în evidență de Gauss și Laplace în studiile lor privind repartiția erorilor.

Importanța fundamentală pe care repartiția normală o are în teoria probabilităților ca și în aplicațiile sale, rezidă, pe de o parte, în faptul că majoritatea fenomenelor supuse la influențe absolut întîmplătoare conduc la funcții de repartiție normale. Pe de altă parte, vom arăta în cele ce urmează că ea este, în anumite condiții, funcția de repartiție a unei sume de variabile aleatoare cu suficient de mulți termeni, putînd fi deci privită ca o *repartiție* sau o *lege limită*.

---

* Vezi § 4.2.

Scanned with OKEN Scanner

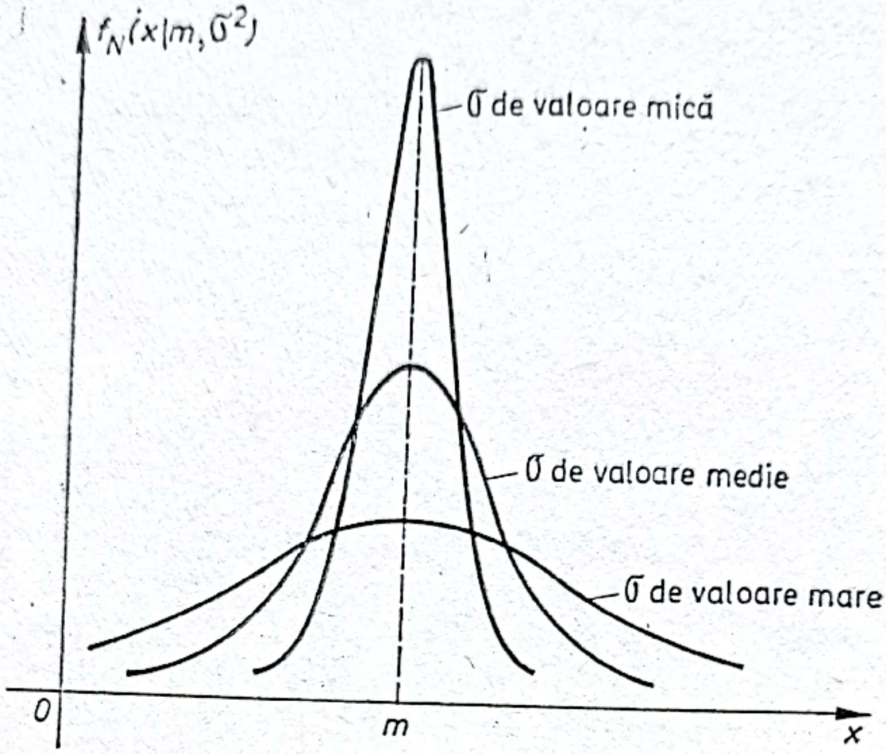


Fig. 10.2.1

10.2. REPARTIȚIA NORMALĂ UNIDIMENSIONALĂ

Repartiția normală este definită de densitatea de probabilitate

$$f_N(x \mid m, \sigma^2) = (2\pi\sigma^2)^{-1/2} \exp\left[-(x-m)^2/2\sigma^2\right], \quad (10.2.1)$$

unde m și $\sigma^2 > 0$ sînt parametri reali. Este clar că funcția f_N este continuă și nenegativă, iar

$$\int_R f_N(x \mid m, \sigma^2)\, dx = 1, \quad (10.2.2)$$

ceea ce justifică faptul că $f_N(x \mid m, \sigma^2)$ este o densitate de probabilitate. Mulțimea variabilelor aleatoare (continue) care au densitatea de probabilitate $f_N(x \mid m, \sigma^2)$ este notată prin $N(m, \sigma^2)$ și este numită *mulțimea variabilelor aleatoare normale* de parametrii m și σ^2.

Graficul densității de probabilitate $f_N(x \mid m, \sigma^2)$ este prezentat în fig. 10.2.1 pentru trei valori diferite ale parametru-

lui σ^2. Graficul are pe $x = m$ ca axă de simetrie, repartiția normală este unimodală (densitatea de probabilitate normală $f_N(x \mid m, \sigma^2)$ are în $x = m$ un punct de maxim, valoarea maximă fiind $f_N(m \mid m, \sigma^2) = (2\pi\sigma^2)^{-1/2}$), iar punctele de abscise $m \pm \sigma$ sînt puncte de inflexiune. Se vede că pentru valori mici ale lui σ^2 curba este „înaltă și ascuțită", iar pentru valori mari ale lui σ^2, curba este „scundă și plată".

Funcția de repartiție corespunzătoare va fi notată prin $F_N(x \mid m, \sigma^2)$ și este dată de relația

$$F_N(x|m, \sigma^2) = (2\pi\sigma^2)^{-1/2} \int_{-\infty}^{x} \exp\ (-(y-m)^2/2\sigma^2)\ \mathrm{d}y. \qquad (10.2.3)$$

Se vede ușor că parametrii m și σ^2 reprezintă respectiv media și dispersia unei variabile aleatoare $\tilde{x} \in N(m, \sigma^2)$. Într-adevăr, făcînd schimbarea de variabilă $t = (y - m)/\sigma$, avem

$$M(\tilde{x}) = (2\pi\sigma^2)^{-1/2} \int_{-\infty}^{\infty} \sigma(m + \sigma t) \mathrm{e}^{-t^2/2}\ \mathrm{d}t = m$$

și

$$V(\tilde{x}) = \sigma^2 (2\pi)^{-1/2} \int_{-\infty}^{\infty} t^2 \mathrm{e}^{-t^2/2}\ \mathrm{d}t = \sigma^2.$$

O variabilă aleatoare normală cu media $m = 0$ și dispersia $\sigma^2 = 1$ este numită variabilă aleatoare normală *standard*. Mulțimea variabilelor aleatoare normale standard este deci notată prin $N(0, 1)$, densitatea de probabilitate corespunzătoare este

$$f_N(x \mid 0, 1) = (2\pi)^{-1/2} \exp\ (-x^2/2), \qquad (10.2.4)$$

iar funcția de repartiție este

$$F_N(x \mid 0, 1) = (2\pi)^{-1/2} \int_{-\infty}^{x} \exp\ (-y^2/2)\ \mathrm{d}y. \qquad (10.2.5)$$

Scanned with OKEN Scanner

Este clar că densitatea de probabilitate $f_N(x \mid 0, 1)$ este o funcție pară, adică

$$f_N(x \mid 0, 1) = f_N(-x \mid 0, 1), \tag{10.2.6}$$

iar funcția de repartiție corespunzătoare $F_N(x \mid 0, 1)$, care se mai notează și cu $\Phi(x)$ și este numită *funcția lui Laplace-Gauss*, are proprietatea

$$F_N(x \mid 0, 1) = 1 - F_N(-x \mid 0, 1). \tag{10.2.7}$$

Din aceste proprietăți reiese că este suficient să cunoaștem $f_N(x \mid 0, 1)$ și $F_N(x \mid 0, 1)$ numai pentru valorile nenegative ale lui x.

Să observăm că schimbarea de variabilă obișnuită în integrala (10.2.3) ne conduce la relația

$$F_N(x \mid m, \sigma^2) = F_N\left(\frac{x-m}{\sigma}\,\middle|\,0, 1\right), \tag{10.2.8}$$

ceea ce reduce calculul lui $F_N(x \mid m, \sigma^2)$ la calculul lui $F_N(x \mid 0, 1)$.

Momente. Vom calcula mai întîi momentele unei variabile aleatoare standard $\tilde{x} \in N(0, 1)$. În acest caz momentele obișnuite și cele centrate coincid.

Momentele de ordin impar sînt

$$M_{2n+1}(\tilde{x}) = V_{2n+1}(\tilde{x}) = (2\pi)^{-1/2} \int_R x^{2n+1} e^{-x^2/2}\, dx = 0, \tag{10.2.9}$$

deoarece funcția care se integrează este impară.

Pentru calculul momentelor de ordin par vom obține o relație de recurență printr-o integrare prin părți:

$$M_{2n}(\tilde{x}) = V_{2n}(\tilde{x}) = (2\pi)^{-1/2} \int_R x^{2n} e^{-x^2/2}\, dx = (2\pi)^{-1/2} \int_R x^{2n-1}\, d(e^{-x^2/2}) =$$

$$= (2\pi)^{-1/2} \left(-x^{2n-1} e^{-x^2/2}\Big|_{-\infty}^{+\infty} + (2n-1) \int_R x^{2n-2} e^{-x^2/2}\, dx\right) = (2n-1) M_{2n-2}(\tilde{x}),$$

de unde se obține relația

$$M_{2n}(\tilde{x}) = V_{2n}(\tilde{x}) = (2n-1)(2n-3)\ldots(3) V_2(\tilde{x}) = 1 \cdot 3 \cdot 5 \ldots (2n-1),$$

dacă ținem seama că $V_2(\tilde{x}) = \sigma^2 = 1$. Rezultatul obținut poate fi scris sub forma

$$M_{2n}(\tilde{x}) = V_{2n}(\tilde{x}) = \frac{(2n)!}{2^n \cdot n!}. \tag{10.2.10}$$

Momentele centrate ale repartiției normale de parametrii m și σ^2 se pot calcula cu schimbarea de variabilă obișnuită, dacă se ține seama de rezultatele (10.2.9) și (10.2.10). Într-adevăr, momentele centrate de ordin impar sînt

$$V_{2n+1}(\tilde{x}) = (2\pi\sigma^2)^{-1/2} \int_R (x-m)^{2n+1} e^{-(x-m)^2/2\sigma^2}\, dx =$$

$$= \sigma^{2n+1}/\sqrt{2\pi} \int_R t^{2n+1} e^{-t^2/2}\, dt = 0, \tag{10.2.11}$$

iar momentele centrate de ordin par se calculează în modul următor:

$$V_{2n}(\tilde{x}) = (2\pi\sigma^2)^{-1/2} \int_R (x-m)^{2n} e^{-(x-m)^2/2\sigma^2}\, dx = \sigma^{2n}/\sqrt{2\pi} \int_R t^{2n} e^{-t^2/2}\, dt$$

și ținînd seama de (10.2.10) obținem

$$V_{2n}(\tilde{x}) = \frac{(2n)!}{2^n \cdot n!} \sigma^{2n} \tag{10.2.12}$$

Momentele obișnuite se obțin cu ajutorul formulelor cunoscute prin utilizarea momentelor centrate (vezi §5.4).

Asimetria și excesul pentru o variabilă aleatoare normală $\tilde{x} \in N(m, \sigma^2)$ sînt respectiv

$$\gamma_1 = V_3(\tilde{x})/(V(\tilde{x}))^{3/2} = 0; \quad \gamma_2 = V_4(\tilde{x})/(V(\tilde{x}))^2 - 3 = 0. \tag{10.2.13}$$

Funcția caracteristică. Să calculăm mai întîi funcția caracteristică a unei variabile normale standard $\tilde{x} \in N(0, 1)$. Avem

$$\varphi_{\tilde{x}}(t) = (2\pi)^{-1/2} \int_R e^{itx} e^{-x^2/2}\, dx = (2\pi)^{-1/2} \int_R \sum_{n=0}^{\infty} \frac{t^n i^n}{n!} x^n e^{-x^2/2}\, dx.$$

Scanned with OKEN Scanner

Deoarece seria din membrul drept este uniform convergentă, putem permuta suma cu integrala și obținem

$$\varphi_{\tilde{x}}(t) = \sum_{n=0}^{\infty} \frac{t^n i^n}{n!} (2\pi)^{-1/2} \int_R x^n e^{-x^2/2}\, dx = \sum_{n=0}^{\infty} \frac{t^n i^n}{n!} V_n(\tilde{x}).$$

Ținînd seama de rezultatele (10.2.9) și (10.2.10) obținem

$$\varphi_{\tilde{x}}(t) = \sum_{m=0}^{\infty} \frac{t^{2m} i^{2m}}{(2m)!} \cdot \frac{(2m)!}{2^m . m!} = \sum_{m=0}^{\infty} \frac{(-t^2/2)^m}{m!} = e^{-t^2/2}. \quad (10.2.14)$$

Să considerăm o variabilă aleatoare $\tilde{x} \in N(m, \sigma^2)$. Variabila aleatoare $\tilde{y} = (\tilde{x} - m)/\sigma$ are evident media zero și dispersia 1 și este de asemenea o variabilă aleatoare normală, adică $\tilde{y} \in N(0, 1)$. Avem deci

$$\varphi_{\tilde{x}}(t) = \varphi_{\sigma\tilde{y}+m}(t) = e^{itm} \varphi_{\tilde{y}}(\sigma t) = e^{itm} e^{-\sigma^2 t^2/2}$$

și prin urmare funcția caracteristică a variabilei $\tilde{x} \in N(m, \sigma^2)$ este

$$\varphi_{\tilde{x}}(t) = \exp(itm - \sigma^2 t^2/2). \quad (10.2.15)$$

În sfîrșit, să observăm că suma unor variabile aleatoare normale independente este tot o variabilă aleatoare normală. Mai exact, avem rezultatul următor.

Teorema 10.2.1. *Dacă* $\tilde{x}_1 \in N(m_1, \sigma_1^2)$ *și* $\tilde{x}_2 \in N(m_2, \sigma_2^2)$ *sînt variabile aleatoare independente, atunci* $\tilde{x} = \tilde{x}_1 + \tilde{x}_2$ *aparține clasei* $N(m, \sigma^2)$, *unde* $m = m_1 + m_2$, *iar* $\sigma^2 = \sigma_1^2 + \sigma_2^2$.

Demonstrație. Funcția caracteristică a sumei unor variabile aleatoare independente este produsul funcțiilor caracteristice. Avem deci

$$\varphi_{\tilde{x}}(t) = \varphi_{\tilde{x}_1}(t) \varphi_{\tilde{x}_2}(t) = \exp(itm_1 - \sigma_1^2 t^2/2) \exp(itm_2 - \sigma_2^2 t^2/2) =$$

$$= \exp(it(m_1 + m_2) - (\sigma_1^2 + \sigma_2^2) t^2/2),$$

care este funcția caracteristică a repartiției normale cu parametrii $m = m_1 + m_2$ și $\sigma^2 = \sigma_1^2 + \sigma_2^2$.

Scanned with OKEN Scanner

10.3. REPARTIȚIA NORMALĂ MULTIDIMENSIONALĂ

Să considerăm vectorul aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_k)'$, vectorul $\mathbf{m} = (m_1, \ldots, m_k)' \in R^k$ și matricea simetrică și pozitiv definită $\mathbf{A} = (a_{ij})_{1 \leqslant i, j \leqslant k}$ cu elemente reale. Spunem că vectorul aleator $\tilde{\mathbf{x}}$ are *repartiție normală k-dimensională (nesingulară)*, dacă densitatea sa de probabilitate este de forma

$$f(\mathbf{x}) = K \exp(-(1/2)(\mathbf{x}' - \mathbf{m}')\mathbf{A}(\mathbf{x} - \mathbf{m})), \qquad (10.3.1)$$

unde $K > 0$ este o constantă ce va fi determinată din condiția ca funcția f definită de (10.3.1), care este evident nenegativă și continuă, să fie o densitate de probabilitate, adică să satisfacă și condiția

$$\int_{R^k} f(x_1, \ldots, x_k)\, dx_1 \ldots dx_k = 1. \qquad (10.3.2)$$

Dintr-o teoremă de algebră liniară știm că orice matrice nesingulară și simetrică poate fi scrisă ca produsul dintre o matrice nesingulară și transpusa acesteia. Deoarece $\mathbf{A}^{-1}$ este nesingulară și simetrică, rezultă că există o matrice $\mathbf{B}$ astfel încît $\mathbf{A}^{-1} = \mathbf{BB}'$ și deci $\mathbf{A} = (\mathbf{BB}')^{-1}$.

Dacă utilizăm matricea $\mathbf{B}$ astfel determinată pentru a face schimbarea de variabile

$$\mathbf{x} = \mathbf{By} + \mathbf{m} \qquad (10.3.3)$$

rezultă că

$$(\mathbf{x} - \mathbf{m})'\mathbf{A}(\mathbf{x} - \mathbf{m}) = (\mathbf{By})'\mathbf{ABy} = \mathbf{y}'\mathbf{B}'\mathbf{ABy} = \mathbf{y}'\mathbf{B}'(\mathbf{BB}')^{-1}\mathbf{By} =$$

$$= \mathbf{y}'\mathbf{B}'(\mathbf{B}')^{-1}\mathbf{B}^{-1}\mathbf{By} = \mathbf{y}'\mathbf{y} = \sum_{j=1}^{k} y_j^2. \qquad (10.3.4)$$

Făcînd schimbarea de variabilă (10.3.3) în integrala (10.3.2) și ținînd seama de (10.3.4), rezultă

$$\int_{R^k} f(x_1, \ldots, x_k)\, dx_1 \ldots, dx_k = \int_{R^k} K e^{-(1/2)\sum_{j=1}^{k} y_j^2} |\det \mathbf{B}|\, dy_1 \ldots, dy_k =$$

$$= K |\det \mathbf{B}| \prod_{j=1}^{k} \int_{R} e^{-y_j^2/2}\, dy_j = K |\det \mathbf{B}| (2\pi)^{k/2}. \qquad (10.3.5)$$

Observînd, în sfîrșit, că $\det(\mathbf{A}^{-1}) = (\det \mathbf{B})(\det \mathbf{B}') = (\det \mathbf{B})^2$, rezultă că

$$\det B = (\det \mathbf{A})^{-1/2} \tag{10.3.6}$$

și deci valoarea căutată a constantei K este

$$K = (2\pi)^{-k/2}(\det \mathbf{A})^{1/2}. \tag{10.3.7}$$

Densitatea de probabilitate a unei repartiții normale (nesingulare) k-dimensionale este deci

$$f(\mathbf{x}) = (2\pi)^{-k/2}(\det \mathbf{A})^{1/2} \exp(-(1/2)(\mathbf{x}-\mathbf{m})'\mathbf{A}(\mathbf{x}-\mathbf{m})). \tag{10.3.8}$$

Să observăm de asemenea că vectorul aleator

$$\tilde{\mathbf{y}} = \mathbf{B}^{-1}(\tilde{\mathbf{x}} - \mathbf{m}) \tag{10.3.9}$$

are densitatea de probabilitate

$$f_{\tilde{\mathbf{y}}}(y_1, \ldots, y_k) = (2\pi)^{-k/2}(\det \mathbf{A})^{1/2} \exp(-(1/2)(\mathbf{By})'\mathbf{A}(\mathbf{By}))|\det \mathbf{B}|,$$

conform formulei (5.1.26). Ținînd seama de (10.3.4) și (10.3.6), rezultă

$$f_{\tilde{\mathbf{y}}}(y_1, \ldots, y_k) = (2\pi)^{-k/2}\exp(-(1/2)\mathbf{y}'\mathbf{y}). \tag{10.3.10}$$

Parametrii $\mathbf{m}$ și $\mathbf{A}$ care apar în densitatea normală k-dimensională au o interpretare interesantă, analoagă celei obținute în cazul unidimensional. Într-adevăr, din relația (10.3.9) rezultă că putem scrie

$$\tilde{\mathbf{x}} = \mathbf{B}\tilde{\mathbf{y}} + \mathbf{m} \tag{10.3.11}$$

și deci media vectorului aleator $\tilde{\mathbf{x}}$ este

$$M(\tilde{\mathbf{x}}) = \mathbf{B}M(\tilde{\mathbf{y}}) + \mathbf{m} \tag{10.3.12}$$

iar matricea de covarianță este

$$V(\tilde{\mathbf{x}}) = M((\tilde{\mathbf{x}} - M(\tilde{\mathbf{x}}))(\tilde{\mathbf{x}} - M(\tilde{\mathbf{x}}))'). \tag{10.3.13}$$

Scanned with OKEN Scanner

Avem însă

$$M(\tilde{y}_j) = \int_{R^k} y_j f\sim(y_1, \ldots, y_k)\, dy_1 \ldots dy_k =$$

$$= \left[(2\pi)^{-1/2} \int_R y_j e^{-y_j^2/2}\, dy_j\right]\left[\prod_{i \neq j} (2\pi)^{-1/2} \int_R e^{-y_i^2/2}\, dy_i\right] = 0,$$

ceea ce arată că $M(\tilde{\mathbf{y}}) = \mathbf{0}$ și deci $M(\tilde{\mathbf{x}}) = \mathbf{m}$, conform relației (10.3.11).

Din (10.3.13) rezultă atunci că

$$V(\tilde{\mathbf{x}}) = M((\tilde{\mathbf{x}} - \mathbf{m})(\tilde{\mathbf{x}} - \mathbf{m})') = M(\mathbf{B}\tilde{\mathbf{y}}(\mathbf{B}\tilde{\mathbf{y}})') =$$

$$= M(\mathbf{B}\tilde{\mathbf{y}}\tilde{\mathbf{y}}'\mathbf{B}') = \mathbf{B}M(\tilde{\mathbf{y}}\tilde{\mathbf{y}}')\mathbf{B}'. \qquad (10.3.14)$$

Pentru $i \neq j$ avem însă

$$M(\tilde{y}_i \tilde{y}_j) = \int_{R^k} y_i y_j f_{\tilde{\mathbf{y}}}(y_1, \ldots, y_k)\, dy_1 \ldots dy_k =$$

$$= \left[(2\pi)^{-1/2} \int_R y_i e^{-y_i^2/2}\, dy_i\right]\left[(2\pi)^{-1/2} \int_R y_j e^{-y_j^2/2}\, dy_j\right] \times$$

$$\times \prod_{\substack{r \neq i \\ r \neq j}} \left[(2\pi)^{-1/2} \int_R y_r e^{-y_r^2/2}\, dy_r\right] = 0,$$

iar pentru $i = j$ avem

$$M(\tilde{y}_i^2) = \int_{R^k} y_i^2 f_{\tilde{\mathbf{y}}}(y_1, \ldots, y_k)\, dy_1 \ldots dy_k =$$

$$= \left[(2\pi)^{-1/2} \int_R y_i^2 e^{-y_i^2/2}\, dy_i\right] \prod_{j \neq i} \left[(2\pi)^{-1/2} \int_R y_j e^{-y_j^2/2}\, dy_j\right] = 1.$$

Scanned with OKEN Scanner

Rezultă atunci că $M(\tilde{\mathbf{y}}\tilde{\mathbf{y}}')$ este matricea unitate și obținem deci

$$V(\tilde{\mathbf{x}}) = \mathbf{B}\mathbf{B}' = \mathbf{A}^{-1}$$

sau

$$\mathbf{A} = (V(\tilde{\mathbf{x}}))^{-1}. \tag{10.3.15}$$

Cu alte cuvinte, parametrii $\mathbf{m}$ și $\mathbf{A}$ din densitatea de probabilitate (10.3.8) reprezintă vectorul valorilor medii $M(\tilde{\mathbf{x}})$, respectiv inversa matricei de covarianță $(V(\tilde{\mathbf{x}}))^{-1}$ a vectorului aleator $\tilde{\mathbf{x}}$.

Dacă notăm, pentru simplitate, matricea de covarianță prin $\boldsymbol{\delta}^2$, adică

$$\boldsymbol{\delta}^2 = V(\tilde{\mathbf{x}}). \tag{10.3.16}$$

iar prin $\boldsymbol{\delta}^{-2}$ inversa matricei de covarianță (sau *matricea de precizie*), adică

$$\boldsymbol{\delta}^{-2} = (V(\tilde{\mathbf{x}}))^{-1} \tag{10.3.17}$$

rezultă că densitatea de probabilitate a repartiției normale (nesingulare) de parametrii $\mathbf{m}$ și $\boldsymbol{\delta}^2$ este

$$f_N(\mathbf{x}|\,\mathbf{m}, \boldsymbol{\delta}^2) = (2\pi)^{-k/2}(\det \boldsymbol{\delta}^2)^{-1/2} \exp\,(-(1/2)(\mathbf{x}-\mathbf{m})'\boldsymbol{\delta}^{-2}(\mathbf{x}-\mathbf{m})). \tag{10.3.18}$$

Mulțimea vectorilor aleatori k-dimensionali care au densitatea de probabilitate $f_N(\mathbf{x} \mid \mathbf{m}, \boldsymbol{\delta}^2)$ va fi notată prin $N_k(\mathbf{m}, \boldsymbol{\delta}^2)$; orice $\tilde{\mathbf{x}} \in N_k(\mathbf{m}, \boldsymbol{\delta}^2)$ va fi numit *vector normal k-dimensional de parametrii* $\mathbf{m}$ și $\boldsymbol{\delta}^2$. În cazul cînd $\mathbf{m} = \mathbf{0}$ și $\boldsymbol{\delta}^2$ este matricea unitate de ordinul k, spunem că vectorul aleator $\tilde{\mathbf{x}} \in N_k(\mathbf{0}, \mathbf{I})$ *are repartiție normală k-dimensională standard*. Densitatea de probabilitate a unui vector aleator $\tilde{\mathbf{x}} \in N_k(\mathbf{0}, \mathbf{I})$ este evident

$$f_N(\mathbf{x} \mid \mathbf{0}, \mathbf{I}) = (2\pi)^{-k/2} \exp\,(-(1/2)\mathbf{x}'\mathbf{x}). \tag{10.3.19}$$

Conform rezultatelor stabilite mai sus, rezultă că orice vector aleator $\tilde{\mathbf{x}} \in N_k(\mathbf{m}, \boldsymbol{\delta}^2)$ se poate obține printr-o transformare liniară nesingulară de forma (10.3.11) dintr-un vector aleator normal standard $\tilde{\mathbf{y}} \in N_k(\mathbf{0}, \mathbf{I})$.

Scanned with OKEN Scanner

În cazul cînd se consideră o transformare liniară de forma

$$\tilde{\mathbf{x}} = \mathbf{B}\tilde{\mathbf{y}} + \mathbf{m} \tag{10.3.20}$$

în care $\tilde{\mathbf{y}} \in N_k(0, \mathbf{I})$, dar $\mathbf{B}$ este o matrice singulară, se obține un vector aleator $\tilde{\mathbf{x}}$ în care cel puțin una dintre componentele sale se poate determina în funcție de celelalte, iar matricea $\boldsymbol{\delta}^2 = \mathbf{B}\mathbf{B}'$ este simetrică, dar numai pozitiv semidefinită. În acest caz vom spune că $\tilde{\mathbf{x}}$ are o *repartiție normală singulară* de parametrii $\mathbf{m}$ și $\boldsymbol{\delta}^2$. Evident, repartiția normală singulară nu are o densitate de probabilitate de forma (10.3.18), deoarece inversa $\boldsymbol{\delta}^{-2}$ nu există.

Teorema 10.3.1. *Fie* $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_k)'$ *vectorul aleator cu repartiție normală singulară de parametrii* $\mathbf{m}$ *și* $\boldsymbol{\delta}^2$ *dat de* (10.3.20). *Dacă* $rang(\boldsymbol{\delta}^2) = \mathrm{r} < \mathrm{k}$, *atunci există o matrice* $\mathbf{C}$ *cu* k *linii și* r *coloane astfel încît*

$$\tilde{\mathbf{x}} = \mathbf{C}\tilde{\mathbf{z}} + \mathbf{m}, \tag{10.3.21}$$

unde $\tilde{\mathbf{z}} \in \mathrm{N}_\mathrm{r}(0, \mathbf{I})$ *și, în plus, avem*

$$\boldsymbol{\delta}^2 = \mathbf{B}\mathbf{B}' = \mathbf{C}\mathbf{C}'. \tag{10.3.22}$$

Demonstrație. Să considerăm $\tilde{\mathbf{x}}$ de forma (10.3.20), unde $\tilde{\mathbf{y}} \in N_k(0, \mathbf{I})$ și $\mathbf{B}$ este singulară. Așa cum am văzut mai sus (vezi (10.3.14)), avem $\mathrm{rang}(\boldsymbol{\delta}^2) = \mathrm{rang}(\mathbf{B}\mathbf{B}')$. Cum $\mathrm{rang}(\boldsymbol{\delta}^2) = r < k$, rezultă $\mathrm{rang}(\mathbf{B}\mathbf{B}') = r$, ceea ce este echivalent cu a spune că $\mathrm{rang}(\mathbf{B}) = r$. După o eventuală renumerotare a componentelor vectorilor $\tilde{\mathbf{x}}$, $\tilde{\mathbf{y}}$ și $\mathbf{m}$ și a liniilor lui $\mathbf{B}$, putem considera că primele r linii din $\mathbf{B}$ sînt liniar independente și formează matricea $\mathbf{B}_1$ cu r linii și k coloane, al cărei rang este r. Așadar, rezultă că putem partiționa $\mathbf{B}$ sub forma

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{pmatrix}$$

și corespunzător

$$\tilde{\mathbf{x}} = \begin{pmatrix} \tilde{\mathbf{x}}_1 \\ \tilde{\mathbf{x}}_2 \end{pmatrix}, \quad \mathbf{m} = \begin{pmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{pmatrix}.$$

Scanned with OKEN Scanner

Deoarece $\text{rang}(B_1) = r$, rezultă că există o matrice D cu $k-r$ linii și k coloane astfel încît $B_2 = DB_1$, deoarece liniile lui B_2 sînt combinații liniare ale liniilor lui B_1. Avem prin urmare

$$\tilde{x} - m = \begin{pmatrix} \tilde{x}_1 - m_1 \\ \tilde{x}_2 - m_2 \end{pmatrix} = \begin{pmatrix} B_1 \\ DB_1 \end{pmatrix} \tilde{y} = \begin{pmatrix} B_1\tilde{y} \\ DB_1\tilde{y} \end{pmatrix},$$

de unde obținem

$$\tilde{x}_2 - m_2 = DB_1\tilde{y} = D(\tilde{x}_1 - m_1)$$

sau

$$\tilde{x}_2 = D\tilde{x}_1 + (m_2 - Dm_1),$$

ceea ce arată că $\tilde{x}_2$ este o funcție liniară de $\tilde{x}_1$. Mai mult, deoarece

$$\tilde{x}_1 = B_1\tilde{y} + m_1,$$

iar $\text{rang}(B_1) = r$, rezultă că $\tilde{x}_1$ este nesingular (vezi § 5.7). Conform cu cele arătate mai sus, există o matrice de ordinul r nesingulară C_1 astfel încît

$$\tilde{x}_1 = C_1\tilde{z} + m_1,$$

unde $\tilde{z} \in N_r(0, I)$.

Dacă notăm $C_2 = DC_1$ și definim

$$C = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix},$$

atunci din faptul că

$$\tilde{x}_2 - m_2 = D(\tilde{x}_1 - m_1) = DC_1\tilde{z} = C_2\tilde{z}$$

obținem

$$\tilde{x} - m = \begin{pmatrix} \tilde{x}_1 - m_1 \\ \tilde{x}_2 - m_2 \end{pmatrix} = \begin{pmatrix} C_1\tilde{z} \\ C_2\tilde{z} \end{pmatrix} = C\tilde{z},$$

ceea ce demonstrează prima afirmație a teoremei.

Pentru a arăta cea de a doua afirmație, să observăm mai întîi că din definiția lui C_1 avem

$$B_1B_1' = C_1C_1'. \qquad (10.3.23)$$

Scanned with OKEN Scanner

Prin urmare

$$\mathbf{BB'} = \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{pmatrix} (\mathbf{B}_1' \quad \mathbf{B}_2') = \begin{pmatrix} \mathbf{B}_1\mathbf{B}_1' & \mathbf{B}_1\mathbf{B}_2' \\ \mathbf{B}_2\mathbf{B}_1' & \mathbf{B}_2\mathbf{B}_2' \end{pmatrix} = \begin{pmatrix} \mathbf{B}_1\mathbf{B}_1' & \mathbf{B}_1\mathbf{B}_1'\mathbf{D}' \\ \mathbf{D}\mathbf{B}_1\mathbf{B}_1' & \mathbf{D}\mathbf{B}_1\mathbf{B}_1'\mathbf{D}' \end{pmatrix}$$

și, ținînd seama de (10.3.23) și de faptul că $\mathbf{C}_2 = \mathbf{D}\mathbf{C}_1$, rezultă

$$\mathbf{BB'} = \begin{pmatrix} \mathbf{C}_1\mathbf{C}_1' & \mathbf{C}_1\mathbf{C}_1'\mathbf{D}' \\ \mathbf{D}\mathbf{C}_1\mathbf{C}_1' & \mathbf{D}\mathbf{C}_1\mathbf{C}_1'\mathbf{D}' \end{pmatrix} = \begin{pmatrix} \mathbf{C}_1\mathbf{C}_1' & \mathbf{C}_1\mathbf{C}_2' \\ \mathbf{C}_2\mathbf{C}_1' & \mathbf{C}_2\mathbf{C}_2' \end{pmatrix} = \begin{pmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \end{pmatrix} (\mathbf{C}_1' \quad \mathbf{C}_2') = \mathbf{CC'}.$$

Observație. Din teorema 10.3.1 rezultă deci că orice vector normal $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_k)'$, singular sau nu, poate fi obținut printr-o transformare liniară dintr-un vector aleator normal standard. Această transformare este de forma (10.3.21), unde $\mathbf{C}$ are k linii și r coloane, iar $r = \operatorname{rang}(\boldsymbol{\delta}^2)$.

Rezultatul obținut ne permite să arătăm că orice vector aleator care se obține printr-o transformare liniară dintr-un vector aleator normal este de asemenea un vector aleator normal. Mai exact, este valabilă teorema următoare.

Teorema 10.3.2. *Fie* $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)$ *un vector aleator cu repartiție normală (nesingulară sau singulară) de parametrii* $\mathbf{m}$ *și* $\boldsymbol{\delta}^2$, *unde* $\operatorname{rang}(\boldsymbol{\delta}^2) = \mathrm{r} \leqslant \mathrm{k}$. *Fie*

$$\tilde{\mathbf{y}} = \mathbf{D}\tilde{\mathbf{x}} + \mathbf{d}, \tag{10.3.24}$$

unde $\mathbf{D}$ *este o matrice cu* p *linii și* k *coloane cu* $\operatorname{rang}(\mathbf{D}) = \mathrm{s}$, *iar* $\mathbf{d} \in \mathrm{R}^{\mathrm{p}}$.

În aceste condiții, $\tilde{}$ *are o repartiție normală de parametrii* $\mathrm{m}_1 = \mathrm{Dm} + \mathrm{d}$ *și* $\delta_1^2 = \mathrm{D}\delta^2\mathrm{D}'$. *Dacă* $\mathrm{p} > \mathrm{k}$, *atunci* $\tilde{\mathbf{y}}$ *este singular. Dacă* $\mathrm{p} \leqslant \mathrm{k}$, $\mathrm{r} = \mathrm{k}$ *și* $\mathrm{s} = \mathrm{p}$, *atunci* $\tilde{}$ *este nesingular.*

Demonstrație. Dacă $r = k$, atunci $\tilde{\mathbf{x}}$ este nesingular și se poate scrie sub forma

$$\tilde{\mathbf{x}} = \mathbf{B}\tilde{\mathbf{u}} + \mathbf{m} \tag{10.3.25}$$

cu $\tilde{\mathbf{u}} \in N_k(\mathbf{0}, \mathbf{I})$. Dacă $r < k$, atunci, conform teoremei precedente, există o matrice $\mathbf{C}$ cu k linii și r coloane astfel încît

$$\mathbf{x} = \mathbf{C}\tilde{\mathbf{v}} + \mathbf{m} \tag{10.3.26}$$

cu $\mathbf{v} \in N_r(\mathbf{0}, \mathbf{I})$.

Scanned with OKEN Scanner

Atunci avem în primul caz

$$\tilde{\mathbf{y}} = \mathbf{D}\tilde{\mathbf{x}} + \mathbf{d} = \mathbf{D}(\mathbf{B}\tilde{\mathbf{u}} + \mathbf{m}) + \mathbf{d} = (\mathbf{DB})\tilde{\mathbf{u}} + (\mathbf{Dm}+\mathbf{d}),$$

unde

$$\operatorname{rang}(\mathbf{DB}) = l \leqslant \min(\operatorname{rang}(\mathbf{D}),\ \operatorname{rang}(\mathbf{B})) = \min(s,\ k).$$

Aplicînd teorema 10.3.1 lui $\tilde{\mathbf{y}}$, rezultă că există o matrice $\mathbf{G}$ cu k linii și l coloane astfel încît să putem scrie

$$\tilde{\mathbf{y}} = \mathbf{G}\tilde{\mathbf{w}} + (\mathbf{Dm} + \mathbf{d}), \tag{10.3.27}$$

unde $\mathbf{w} \in N_l(\mathbf{0}, \mathbf{I})$. Rezultă atunci că $\tilde{\mathbf{y}}$ are o repartiție normală cu parametrii $\mathbf{m}_1 = \mathbf{Dm} + \mathbf{d}$ și $\boldsymbol{\delta}_1^2 = \mathbf{GG}'$. Din (10.3.22) reiese însă că

$$\mathbf{GG}' = \mathbf{DB}(\mathbf{DB})' = \mathbf{DBB}'\mathbf{D}' = \mathbf{D}\boldsymbol{\delta}^2\mathbf{D}',$$

ceea ce demonstrează afirmația în acest caz.

În cazul cînd $\tilde{\mathbf{x}}$ este de forma (10.3.26), demonstrația se desfășoară în același mod. Într-adevăr,

$$\tilde{\mathbf{y}} = \mathbf{D}\tilde{\mathbf{x}} + \mathbf{d} = \mathbf{D}(\mathbf{C}\tilde{\mathbf{v}} + \mathbf{m}) + \mathbf{d} = (\mathbf{DC})\tilde{\mathbf{v}} + (\mathbf{Dm} + \mathbf{d}),$$

unde

$$\operatorname{rang}(\mathbf{DC}) = l' \leqslant \min(\operatorname{rang}(\mathbf{D}),\ \operatorname{rang}(\mathbf{C})) = \min(s, r).$$

Aplicînd teorema 10.3.1 lui $\tilde{\mathbf{y}}$, rezultă că există o matrice $\mathbf{H}$ cu k linii și l' coloane astfel încît să putem scrie

$$\tilde{\mathbf{y}} = \mathbf{H}\tilde{\mathbf{z}} + (\mathbf{Dm} + \mathbf{d}), \tag{10.3.28}$$

unde $\tilde{\mathbf{z}} \in N_{l'}(\mathbf{0}, \mathbf{I})$. Rezultă atunci că $\tilde{\mathbf{y}}$ are o repartiție normală cu parametrii $\mathbf{m}_1 = \mathbf{Dm}+\mathbf{d}$ și $\boldsymbol{\delta}_1^2 = \mathbf{HH}'$. Din (10.3.28) avem însă

$$\mathbf{HH}' = \mathbf{DC}(\mathbf{DC})' = \mathbf{DCC}'\mathbf{D}' = \mathbf{D}\boldsymbol{\delta}^2\mathbf{D}',$$

ceea ce încheie demonstrația primei afirmații.

A doua și a treia afirmație rezultă imediat din (10.3.27) și (10.3.28), dacă se ține seama că rangul produsului a două matrice este mai mic decît fiecare dintre rangurile factorilor.

Scanned with OKEN Scanner

Observație. Să considerăm cazul particular în care vectorul aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_k)'$ are componente independente și identic repartizate din clasa $N(0, 1)$. Atunci, evident, $\boldsymbol{\delta}^2 = \mathbf{I}$ și $\mathbf{m} = \mathbf{0}$. Dacă $\mathbf{D}$ este o matrice ortonormală de ordinul k, rezultă imediat că $\tilde{\mathbf{y}}$ dat de (10.3.24) este de asemenea din clasa $N_k(\mathbf{0}, \mathbf{I})$, ceea ce arată că $\tilde{\mathbf{y}}$ are componente independente și identic repartizate din clasa $N(0, 1)$.

Utilizînd teorema precedentă, putem determina ușor repartiția unui vector format cu o parte a componentelor unui vector aleator normal. Mai precis, avem rezultatul următor.

Teorema 10.3.3. *Dacă* $\tilde{\mathbf{x}} = (\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_k)'$ *are repartiție normală de parametrii* $\mathbf{m}$ *și* $\boldsymbol{\delta}^2$, *iar* $\tilde{\mathbf{x}}_1 = (\tilde{\mathbf{x}}_{i_1}, \ldots, \tilde{\mathbf{x}}_{i_p})$, *atunci* $\tilde{\mathbf{x}}_1$ *are repartiție normală de parametrii* $\mathbf{m}_1$ și $\boldsymbol{\delta}_{11}^2$, *unde*

$$\mathbf{m}_1 = (\mathbf{m}_{i_1}, \ldots, \mathbf{m}_{i_p})'$$

și

$$\boldsymbol{\delta}_{11}^2 = \mathrm{V}(\tilde{\mathbf{x}}_{i_1}, \ldots, \tilde{\mathbf{x}}_{i_p}).$$

Demonstrație. Fie $\mathbf{D}_1$ matricea obținută din matricea unitate de ordinul k prin eliminarea liniilor diferite de cele care au rangurile $i_1, \ldots, i_p$. Este clar că putem scrie

$$\tilde{\mathbf{x}}_1 = \mathbf{D}_1 \tilde{\mathbf{x}} \tag{10.3.29}$$

Aplicînd teorema 10.3.2 vectorului aleator $\tilde{\mathbf{x}}_1$ dat de relația (10.3.29) rezultă că $\tilde{\mathbf{x}}_1$ are o repartiție normală de parametrii $\mathbf{D}_1\mathbf{m} + \mathbf{0} = \mathbf{m}_1$ și $\mathbf{D}_1\boldsymbol{\delta}^2\mathbf{D}_1' = \boldsymbol{\delta}_{11}^2$.

Observație. Este clar că dacă $\tilde{\mathbf{x}}$ este nesingular, atunci $\tilde{\mathbf{x}}_1$ este de asemenea nesingular. Cu alte cuvinte, dacă $\tilde{\mathbf{x}} \in N_k(\mathbf{m}, \boldsymbol{\delta}^2)$, are densitatea de probabilitate $f_N(\mathbf{x} \mid \mathbf{m}, \boldsymbol{\delta}^2)$, atunci $\tilde{\mathbf{x}}_1$ are densitatea de probabilitate $f_N(\mathbf{x}_1 \mid \mathbf{m}_1, \boldsymbol{\delta}_{11}^2)$.

De remarcat însă că dacă $\tilde{\mathbf{x}}_1$ este nesingular, nu rezultă în mod necesar că $\tilde{\mathbf{x}}$ este nesingular. Într-adevăr, fie $\tilde{\mathbf{y}} \in N_2(\mathbf{0}, \mathbf{I})$ și fie $\tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2)'$ vectorul aleator definit de relația

$$\tilde{\mathbf{x}} = \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{pmatrix}.$$

Este clar că $\tilde{x}_1 = \tilde{y}_1$ și $\tilde{x}_2 = \tilde{y}_2$ sînt nesingulare, dar vectorul $\tilde{\mathbf{x}}$ este singular (matricea transformării are rangul egal cu 1).

Scanned with OKEN Scanner

Vom calcula acum funcția caracteristică a unui vector aleator normal $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_k)'$ de parametrii $\mathbf{m}$ și $\boldsymbol{\delta}^2$. Așa cum am văzut mai sus, există o matrice $\mathbf{C}$ cu k linii și r coloane astfel încît să avem

$$\tilde{\mathbf{x}} = \mathbf{C}\tilde{\mathbf{z}} + \mathbf{m},$$

unde $\tilde{\mathbf{z}} \in N_r(\mathbf{0}, \mathbf{I})$, iar $r = \operatorname{rang}(\boldsymbol{\delta}^2)$. În plus, $\boldsymbol{\delta}^2 = \mathbf{CC}'$. Cu aceste observații, rezultă că funcția caracteristică a vectorului aleator $\tilde{\mathbf{x}}$ este

$$\varphi_{\tilde{\mathbf{x}}}(t) = M(e^{-t'\tilde{\mathbf{x}}}) = M(e^{it'\mathbf{C}\tilde{\mathbf{z}}+it'\mathbf{m}}) = e^{it'\mathbf{m}} M(e^{it'\mathbf{C}\tilde{\mathbf{z}}}). \quad (10.3.30)$$

Dacă notăm prin c_{js}, $1 \leqslant j \leqslant k$, $1 \leqslant s \leqslant r$, elementele matricei $\mathbf{C}$, rezultă că putem scrie

$$t'\mathbf{C}\tilde{\mathbf{z}} = \sum_{j=1}^{k} \sum_{s=1}^{r} t_j c_{js} \tilde{z}_s = \sum_{s=1}^{r} u_s \tilde{z}_s, \quad (10.3.31)$$

unde

$$u_s = \sum_{j=1}^{k} t_j c_{js}. \quad (10.3.32)$$

Din (10.3.30) și (10.3.31) și din faptul că $\tilde{z}_1, \ldots, \tilde{z}_r$ sînt variabile aleatoare normale standard independente, obținem

$$\varphi_{\mathbf{x}}(t) = e^{it'\mathbf{m}} M\left(e^{i\sum_{s=1}^{r} u_s \tilde{z}_s}\right) = e^{it'\mathbf{m}} \prod_{s=1}^{r} M(e^{iu_s\tilde{z}_s}) = e^{it'\mathbf{m}} \prod_{s=1}^{r} \varphi_{\tilde{z}_s}(u_s) =$$

$$= e^{it'\mathbf{m}} \prod_{s=1}^{r} e^{-u_s^2/2} = e^{it'\mathbf{m}} e^{(-1/2)\sum_{s=1}^{r} u_s^2}. \quad (10.3.33)$$

Avem însă

$$\sum_{s=1}^{r} u_s^2 = \sum_{s=1}^{r} \sum_{j,l=1}^{k} t_j c_{jr} t_l c_{lr} = \sum_{j,l=1}^{k} t_j t_l \sum_{s=1}^{r} c_{js} c_{ls} = t'\mathbf{CC}'t = t'\boldsymbol{\delta}^2 t. \quad (10.3.34)$$

Din (10.3.33) și (10.3.34) se deduce atunci imediat

$$\varphi_{\tilde{\mathbf{x}}}(t) = e^{it'\mathbf{m} - (1/2)t'\boldsymbol{\delta}^2 t}. \quad (10.3.35)$$

Evident, funcția caracteristică a unui vector aleator $\tilde{\mathbf{x}} \in N_k(\mathbf{0}, \mathbf{I})$ este

$$\varphi_{\tilde{\mathbf{x}}}(t) = e^{-(1/2)t't} = e^{-(1/2)\sum_{j=1}^{k} t_j^2}. \quad (10.3.36)$$

Teorema 10.3.4. *Să considerăm vectorul aleator normal* $\tilde{\mathbf{x}}_1$ *cu parametrii* $\mathbf{m}_1$ *și* $\boldsymbol{\delta}_1^2$ *și vectorul aleator normal* $\tilde{\mathbf{x}}_2$ *cu parametrii* $\mathbf{m}_2$ *și* $\boldsymbol{\delta}_2^2$. *Dacă* $\tilde{\mathbf{x}}_1$ *și* $\tilde{\mathbf{x}}_2$ *sînt independenți, atunci vectorul aleator* $\tilde{\mathbf{x}} = \tilde{\mathbf{x}}_1 + \tilde{\mathbf{x}}_2$ *este de asemenea normal cu parametrii* $\mathbf{m} = \mathbf{m}_1 + \mathbf{m}_2$ *și* $\boldsymbol{\delta}^2 = \boldsymbol{\delta}_1^2 + \boldsymbol{\delta}_2^2$. *Dacă* $\tilde{\mathbf{x}}_1$ *și* $\tilde{\mathbf{x}}_2$ *sînt nesingulari, atunci* $\tilde{\mathbf{x}}$ *este de asemenea nesingular.*

Demonstrație. Rezultatul se obține imediat, folosind funcția caracteristică:

$$\varphi_{\tilde{\mathbf{x}}}(\mathbf{t}) = \varphi_{\tilde{\mathbf{x}}_1}(\mathbf{t}) \cdot \varphi_{\tilde{\mathbf{x}}_2}(\mathbf{t}) = e^{i\mathbf{t}'\mathbf{m}_1 - (1/2)\mathbf{t}'\boldsymbol{\delta}_1^2\mathbf{t}} \cdot e^{i\mathbf{t}'\mathbf{m}_2 - (1/2)\mathbf{t}'\boldsymbol{\delta}_2^2\mathbf{t}} =$$

$$= e^{i\mathbf{t}'(\mathbf{m}_1+\mathbf{m}_2) - (1/2)\mathbf{t}'(\boldsymbol{\delta}_1^2+\boldsymbol{\delta}_2^2)\mathbf{t}}.$$

Cu ajutorul funcției caracteristice putem calcula momentele de diverse ordine ale unui vector aleator; evident, tot pe această cale se pot reinterpreta parametrii $\mathbf{m}$ și $\boldsymbol{\delta}^2$ ai repartiției normale multidimensionale.

Să considerăm vectorul aleator normal nesingular $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_k)'$ cu parametrii $\mathbf{m}$ și $\boldsymbol{\delta}^2$. Fie $\tilde{\mathbf{x}}_1 = (\tilde{x}_1, \ldots, \tilde{x}_r)'$ și $\tilde{\mathbf{x}}_2 = (\tilde{x}_{r+1}, \ldots, \tilde{x}_k)'$ vectori aleatori obținuți prin partiționarea lui $\tilde{\mathbf{x}}$. Vom partiționa corespunzător vectorul $\mathbf{m}$ și matricea $\boldsymbol{\delta}^2$:

$$\mathbf{m} = \begin{pmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{pmatrix}, \quad \boldsymbol{\delta}^2 = \begin{pmatrix} \boldsymbol{\delta}_{11}^2 & \boldsymbol{\delta}_{12}^2 \\ \boldsymbol{\delta}_{21}^2 & \boldsymbol{\delta}_{22}^2 \end{pmatrix}. \tag{10.3.37}$$

Teorema 10.3.5. *Dacă* $\tilde{\mathbf{x}} \in \mathrm{N}_k(\mathbf{m}, \boldsymbol{\delta}^2)$ *este nesingular, atunci densitatea de probabilitate a lui* $\tilde{\mathbf{x}}_2$ *condiționată de* $\tilde{\mathbf{x}}_1 = \mathbf{x}_1$ *este o densitate de probabilitate normală* (k—r)*—dimensională cu parametrii* $\mathbf{m}_{2|1}(\mathbf{x}_1)$ *și* $\boldsymbol{\delta}_{22|1}^2$, *unde*

$$\mathbf{m}_{2|1}(\mathbf{x}_1) = \mathbf{m}_2 + \boldsymbol{\delta}_{21}^2(\boldsymbol{\delta}_{11}^2)^{-1}(\mathbf{x}_1 - \mathbf{m}_1) \tag{10.3.38}$$

și

$$\boldsymbol{\delta}_{22|1}^2 = \boldsymbol{\delta}_{22}^2 - \boldsymbol{\delta}_{21}^2(\boldsymbol{\delta}_{11}^2)^{-1}\boldsymbol{\delta}_{12}^2. \tag{10.3.39}$$

Demonstrație. Deoarece $\tilde{\mathbf{x}}$ este nesingular, rezultă că matricea inversă $\boldsymbol{\delta}^{-2} = (\boldsymbol{\delta}^2)^{-1}$ există. Partiționăm această matrice în conformitate cu partiția lui $\boldsymbol{\delta}^2$, punînd

$$\boldsymbol{\delta}^{-2} = \begin{pmatrix} \boldsymbol{\alpha}_{11} & \boldsymbol{\alpha}_{12} \\ \boldsymbol{\alpha}_{21} & \boldsymbol{\alpha}_{22} \end{pmatrix}. \tag{10.3.40}$$

Scanned with OKEN Scanner

Rezultă imediat că

$$\begin{aligned} &\alpha_{11}\delta_{11}^2 + \alpha_{12}\delta_{21}^2 = \mathbf{I}_r, \\ &\alpha_{21}\delta_{12}^2 + \alpha_{22}\delta_{22}^2 = \mathbf{I}_{k-r}, \\ &\alpha_{11}\delta_{12}^2 + \alpha_{12}\delta_{22}^2 = 0, \\ &\alpha_{21}\delta_{11}^2 + \alpha_{22}\delta_{21}^2 = 0, \end{aligned} \tag{10.3.41}$$

unde $\mathbf{I}_s$ reprezintă matricea unitate de ordinul s.

Din ultima relație din (10.3.41) obținem

$$\delta_{21}^2 = -\alpha_{22}^{-1}\alpha_{21}\delta_{11}^2,$$

care, împreună cu prima relație din (10.3.41), conduce la

$$(\delta_{11}^2)^{-1} = \alpha_{11} - \alpha_{12}\alpha_{22}^{-1}\alpha_{21}. \tag{10.3.42}$$

Din a doua și a treia relație din (10.3.41) se obține

$$\alpha_{22} = (\delta_{22}^2 - \delta_{21}^2(\delta_{11}^2)^{-1}\delta_{12}^2)^{-1}. \tag{10.3.43}$$

În sfîrșit, înmulțind ultima relație din (10.3.41) cu α_{22}^{-1} obținem

$$\alpha_{22}^{-1}\alpha_{21} = -\delta_{21}^2(\delta_{11}^2)^{-1}. \tag{10.3.44}$$

Deoarece $\tilde{\mathbf{x}} \in N_k(\mathbf{m}, \delta^2)$, rezultă că densitatea sa de probabilitate este

$$f_N(\mathbf{x}|\mathbf{m}, \delta^2) = (2\pi)^{-k/2}(\det \delta^2)^{-1/2}\exp\left(-(1/2)Q(\mathbf{x})\right),$$

unde

$$Q(\mathbf{x}) = (\mathbf{x}-\mathbf{m})'\delta^{-2}(\mathbf{x}-\mathbf{m}) = \begin{pmatrix} \mathbf{x}_1 - \mathbf{m}_1 \\ \mathbf{x}_2 - \mathbf{m}_2 \end{pmatrix}' \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 - \mathbf{m}_1 \\ \mathbf{x}_2 - \mathbf{m}_2 \end{pmatrix}.$$

Dacă notăm

$$\mathbf{y}_1 = \mathbf{x}_1 - \mathbf{m}_1, \quad \mathbf{y}_2 = \mathbf{x}_2 - \mathbf{m}_2,$$

atunci putem scrie

$$Q(\mathbf{x}) = \mathbf{y}_1'\alpha_{11}\mathbf{y}_1 + \mathbf{y}_1'\alpha_{12}\mathbf{y}_2 + \mathbf{y}_2'\alpha_{21}\mathbf{y}_1 + \mathbf{y}_2'\alpha_{22}\mathbf{y}_2. \tag{10.3.45}$$

Scanned with OKEN Scanner

Analog, deoarece, conform teoremei 10.3.3, avem $\tilde{\mathbf{x}}_1 \in N_r(\mathbf{m}_1, \delta_{11}^2)$, rezultă că densitatea de probabilitate a lui $\tilde{\mathbf{x}}_1$ este

$$f_N(\mathbf{x}_1|\mathbf{m}_1, \delta_{11}^2) = (2\pi)^{-r/2}(\det \delta_{11}^2)^{-1/2}\exp\left((-1/2)Q_1(\mathbf{x}_1)\right),$$

unde

$$Q_1(\mathbf{x}_1)=(\mathbf{x}_1-\mathbf{m}_1)'(\delta_{11}^2)^{-1}(\mathbf{x}_1-\mathbf{m}_1)=\mathbf{y}_1'(\delta_{11}^2)^{-1}\mathbf{y}_1=$$
$$=\mathbf{y}_1'(\alpha_{11}-\alpha_{12}\alpha_{22}^{-1}\alpha_{21})\mathbf{y}_1, \qquad (10.3.46)$$

dacă ținem seama și de (10.3.42). Rezultă atunci că densitatea de probabilitate a lui $\tilde{\mathbf{x}}_2$ condiționată de $\tilde{\mathbf{x}}_1 = \mathbf{x}_1$ este

$$f(\mathbf{x}_2|\mathbf{x}_1) = \frac{f_N(\mathbf{x}|\mathbf{m}, \delta^2)}{f_N(\mathbf{x}_1|\mathbf{m}_1, \delta_{11}^2)} =$$
$$= (2\pi)^{-(k-r)/2}[(\det \delta^2)/(\det \delta_{11}^2)]^{-1/2}\exp\left(-(1/2)Q_{21}\right), \qquad (10.3.47)$$

unde

$$Q_{21} = Q(\mathbf{x}) - Q(\mathbf{x}_1).$$

Din relațiile (10.3.45) și (10.3.46) rezultă însă

$$Q_{21} = \mathbf{y}_1'\alpha_{12}\mathbf{y}_2 + \mathbf{y}_2'\alpha_{21}\mathbf{y}_1 + \mathbf{y}_2'\alpha_{22}\mathbf{y}_2 + \mathbf{y}_1'\alpha_{12}\alpha_{22}^{-1}\alpha_{21}\mathbf{y}_1.$$

Ținînd seama că matricele α, α_{22} și α_{22}^{-1} sînt simetrice, rezultă că putem scrie

$$Q_{21} = \mathbf{y}_2'\alpha_{22}\mathbf{y}_2 + (\alpha_{22}^{-1}\alpha_{21}\mathbf{y}_1)'\alpha_{22}\mathbf{y}_2 + \mathbf{y}_2'\alpha_{22}\alpha_{22}^{-1}\alpha_{21}\mathbf{y}_1 +$$
$$+ (\alpha_{22}^{-1}\alpha_{21}\mathbf{y}_1)'\alpha_{22}(\alpha_{22}^{-1}\alpha_{21}\mathbf{y}_1)$$

și, după ce se dau factorii comuni între primii doi și ultimii doi termeni, obținem

$$Q_{21} = (\mathbf{y}_2 + \alpha_{22}^{-1}\alpha_{21}\mathbf{y}_1)'\alpha_{22}(\mathbf{y}_2 + \alpha_{22}^{-1}\alpha_{21}\mathbf{y}_1).$$

Ținînd seama de (10.3.39), (10.3.43) și (10.3.44) se poate scrie

$$Q_{21} = (\mathbf{x}_2 - \mathbf{m}_2 - \delta_{21}^2(\delta_{11}^2)^{-1}(\mathbf{x}_1 - \mathbf{m}_1))'(\delta_{22/1}^2)^{-1}(\mathbf{x}_2 - \mathbf{m}_2 -$$
$$- \delta_{21}^2(\delta_{11}^2)^{-1}(\mathbf{x}_1 - \mathbf{m}_1))$$

sau încă, avînd în vedere și (10.3.38), rezultă

$$Q_{21} = (\mathbf{x}_2 - \mathbf{m}_{2/1}(\mathbf{x}_1))'(\delta_{22/1}^2)^{-1}(\mathbf{x}_2 - \mathbf{m}_{2/1}(\mathbf{x}_1))'. \qquad (10.3.48)$$

Scanned with OKEN Scanner

În sfîrșit, se observă ușor că

$$(\det(\delta^2_{22/1}))(\det(\delta^2_{11})) = \det(\delta^2). \tag{10.3.49}$$

Într-adevăr, ținînd seama de (10.3.43), rezultă

$$\begin{pmatrix} \delta^2_{11} & \delta^2_{12} \\ \delta^2_{21} & \delta^2_{22} \end{pmatrix} \begin{pmatrix} \mathbf{I}_r & -(\delta^2_{11})^{-1}\delta^2_{12}a_{22} \\ \mathbf{0} & a_{22} \end{pmatrix} = \begin{pmatrix} \delta^2_{11} & \mathbf{0} \\ \delta^2_{21} & \mathbf{I}_{k-r} \end{pmatrix},$$

de unde

$$(\det(\delta^2))\,(\det(a_{22})) = \det(\delta^2_{11})$$

și deci, conform (10.3.39) și (10.3.43) obținem (10.3.49).

Este evident acum că din (10.3.47), (10.3.48) și (10.3.49) rezultă

$$f(\mathbf{x}_2|\mathbf{x}_1) = f_N(\mathbf{x}_2|\mathbf{m}_{2/1}(\mathbf{x}_1),\ \delta^2_{22/1}), \tag{10.3.50}$$

ceea ce dealtfel trebuia arătat.

Observație. Densitatea de probabilitate a lui $\tilde{\mathbf{x}}_2$ condiționată de $\tilde{\mathbf{x}}_1 = \mathbf{x}_1$ este deci o densitate de probabilitate normală $(k-r)$-dimensională cu parametrii $\mathbf{m}_{2/1}(\mathbf{x}_1)$ și $\delta^2_{22/1}$. Media condiționată a lui $\tilde{\mathbf{x}}_2$ este deci o funcție liniară de $\mathbf{x}_1$, iar matricea de covarianță condiționată $\delta^2_{22/1}$ nu depinde de $\mathbf{x}_1$, ceea ce are o mare importanță în aplicațiile practice.

Vom da acum o generalizare a teoremei (10.3.2), utilă în aplicații.

Teorema 10.3.6. *Fie* $\tilde{\mathbf{x}} = (\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_{kp})'$ *un vector aleator normal cu parametrii* $\mathbf{m}$ *și* δ^2. *Să partiționăm* $\tilde{\mathbf{x}}$ *în vectorii* $\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_p$, *unde* $\tilde{\mathbf{x}}_j = (\tilde{\mathbf{x}}_{k(j-1)+1}, \ldots, \tilde{\mathbf{x}}_{kj})$, $1 \leqslant j \leqslant p$. *Partiționăm corespunzător* $\mathbf{m}$ *în vectorii* $\mathbf{m}_1, \ldots, \mathbf{m}_p$ *și matricea* δ^2 *în matricele* δ^2_{ij}, $1 \leqslant i, j \leqslant p$.

Fie $\mathbf{D}_{rj}$, $1 \leqslant r \leqslant s$, $1 \leqslant j \leqslant p$, *o matrice cu* l *linii și* k *coloane și* $\mathbf{d}_r$, $1 \leqslant r \leqslant s$, *vectori din* R^l. *Dacă se definesc vectorii aleatori* $\tilde{\mathbf{y}}_r$, $1 \leqslant r \leqslant s$, *prin*

$$\tilde{\mathbf{y}}_r = \sum_{j=1}^{p} \mathbf{D}_{rj}\tilde{\mathbf{x}}_j + \mathbf{d}_r, \tag{10.3.51}$$

atunci vectorul $\tilde{y} = (\tilde{y}_1', \ldots, \tilde{y}_s')'$ *cu* sl *componente este un vector aleator normal de parametrii* $\overline{m} = (\overline{m}_1', \ldots, \overline{m}_s')'$ *și*

$$\overline{\delta}^2 = \begin{pmatrix} \overline{\delta}_{11}^2 \cdots \overline{\delta}_{1s}^2 \\ \overline{\delta}_{s1}^2 \cdots \overline{\delta}_{ss}^2 \end{pmatrix}, \tag{10.3.52}$$

unde

$$\overline{m}_r = \sum_{j=1}^{p} D_{rj} m_j + d_r, \quad 1 \leqslant r \leqslant s, \tag{10.3.53}$$

$$\overline{\delta}_{rh}^2 = \sum_{i=1}^{p} \sum_{j=1}^{p} D_{ri} \delta_{ij}^2 D_{hj}', \quad 1 \leqslant r, h \leqslant s. \tag{10.3.54}$$

Demonstrație Să considerăm vectorul cu *sl* componente $d = (d_1', \ldots, d_s')'$ și matricea $D = (D_{rj})$, $1 \leqslant r \leqslant s$, $1 \leqslant j \leqslant p$ cu *sl* linii și *kp* coloane. Din definiția lui $\tilde{x}$ și $\tilde{y}$ este clar că

$$\tilde{y} = D\tilde{x} + d. \tag{10.3.55}$$

Conform teoremei 10.3.2, rezultă că $\tilde{y}$ este un vector aleator normal cu parametrii $\overline{m} = Dm + d$ și $\delta^2 = D\delta^2 D'$.

Utilizînd partițiile introduse mai sus pentru vectorii și matricele considerate rezultă imediat relațiile (10.5.53) și (10.5.54) și prin urmare teorema este demonstrată.

Observație. În cazul particular $s = l = k = 1$ rezultă că $\tilde{x}$ este un vector aleator normal cu p componente scalare, avînd parametrii m și δ^2, iar $\tilde{y}$ este de forma

$$\tilde{y} = a'\tilde{x} + b,$$

cu $a \in R^p$ și $b \in R$. Din rezultatul obținut mai sus rezultă că $\tilde{y}$ este o variabilă aleatoare normală de parametrii $\overline{m} = a'm + b$ și $\overline{\sigma}^2 = a'\delta^2 a$.

O altă teoremă utilă în aplicații este teorema următoare.

Teorema 10.3.7. *Fie* $\tilde{x}_i$, $1 \leqslant i \leqslant p$, *vectori aleatori independenți din clasa* $N_k(m_i, \delta^2)$, $1 \leqslant i \leqslant p$. *Dacă* $D = (d_{ij})$ *este o matrice ortogonală de ordinul* p, *atunci vectorii aleatori*

$$\tilde{y}_i = \sum_{j=1}^{p} d_{ij} \tilde{x}_j, \quad 1 \leqslant i \leqslant p,$$

Scanned with OKEN Scanner

sînt de asemenea independenți și aparțin claselor $N(\overline{m}_i, \delta^2)$, $1 \leqslant i \leqslant p$, *unde*

$$\overline{m}_i = \sum_{j=1}^{p} d_{ij} m_j, \quad 1 \leqslant i \leqslant p.$$

Demonstrație. Componentele vectorilor aleatori $\tilde{y}_i$ sînt combinații liniare de variabile aleatoare normale și deci, conform teoremei precedente, sînt variabile aleatoare normale. Mai mult, avem

$$M(\tilde{y}_i) = \sum_{j=1}^{p} d_{ij} M(\tilde{x}_j) = \sum_{j=1}^{p} d_{ij} m_j$$

și

$$\operatorname{cov}(\tilde{y}_i, \tilde{y}_j) = \operatorname{cov}\left(\sum_{r=1}^{p} d_{ir}\tilde{x}_r, \sum_{s=1}^{p} d_{js}\tilde{x}_s\right) =$$

$$= \sum_{r=1}^{p}\sum_{s=1}^{p} d_{ir} d_{js} \operatorname{cov}(\tilde{x}_r, \tilde{x}_s) = \sum_{r=1}^{p}\sum_{s=1}^{p} d_{ir} d_{js} \delta_{rs} \delta^2$$

sau

$$\operatorname{cov}(\tilde{y}_i, \tilde{y}_j) = \delta_{ij}\delta^2,$$

unde δ_{ij} și δ_{rs} sînt simbolurile lui Krönecker. Din ultima relație se deduc imediat celelalte afirmații ale teoremei.

Vom încheia prezentarea repartiției normale multidimensionale făcînd cîteva precizări pentru cazul bidimensional.

Dacă $\tilde{x} = (\tilde{x}_1, \tilde{x}_2)'$ este repartizat normal cu parametrii m și δ^2, atunci

$$m = (m_1, m_2)'$$

și

$$\delta^2 = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 r \\ \sigma_1\sigma_2 r & \sigma_2^2 \end{pmatrix},$$

unde

$$\sigma_1^2 = V(\tilde{x}_1), \quad \sigma_2^2 = V(\tilde{x}_2), \quad r = r(\tilde{x}_1, \tilde{x}_2) = \frac{\operatorname{cov}(\tilde{x}_1, \tilde{x}_2)}{\sigma_1\sigma_2}.$$

Scanned with OKEN Scanner

Rezultă

$$\det(\delta^2) = \sigma_1^2\sigma_2^2(1-r^2)$$

și deci

$$(\delta^2)^{-1} = \frac{1}{1-r^2}\begin{pmatrix} 1/\sigma_1^2 & -r/\sigma_1\sigma_2 \\ -r/\sigma_1\sigma_2 & 1/\sigma_2^2 \end{pmatrix}. \qquad (10.3.56)$$

Densitatea de probabilitate corespunzătoare este prin urmare

$$f_N(x_1, x_2|\mathbf{m}, \delta^2) = (2\pi\sigma_1\sigma_2)^{-1}(1-r^2)^{-1/2}\exp\left(-\frac{1}{2(1-r^2)}\left[\frac{(x_1-m_1)^2}{\sigma_1^2} - \right.\right.$$

$$\left.\left. - \frac{2r(x_1-m_1)(x_2-m_2)}{\sigma_1\sigma_2} + \frac{(x_2-m_2)^2}{\sigma_2^2}\right]\right). \qquad (10.3.57)$$

În sfîrșit, pentru a determina densitatea de probabilitate condiționată $f(x_2|x_1)$, să observăm mai întîi că

$$m_{2/1}(x_1) = m_2 + \sigma_1\sigma_2 r(x_1-m_1)/\sigma_1^2 = m_2 + \sigma_2 r(x_1-m_1)/\sigma_1 \qquad (10.3.58)$$

și

$$\delta_{22/1}^2 = \sigma_2^2 - (\sigma_1\sigma_2 r)^2/\sigma_1^2 = \sigma_2^2(1-r^2). \qquad (10.3.59)$$

Obținem deci

$$f(x_2 \mid x_1) = \frac{1}{\sigma_2\sqrt{2\pi(1-r^2)}}\exp\left[-\frac{1}{2}\left(\frac{x_2 - m_2 - r\sigma_2(x_1-m_1)/\sigma_1}{\sigma_2\sqrt{1-r^2}}\right)^2\right]. \qquad (10.3.60)$$

Din formula (10.3.59) putem spune că r^2 reprezintă proporția în care se reduce σ_2^2 din cauza informației suplimentare $\widetilde{x}_1 = x_1$.

10.4. PROBLEMA ASIMPTOTICĂ CENTRALĂ

Să considerăm un șir de variabile aleatoare independente $(\tilde{x}_n)$ pe cîmpul borelian de probabilitate $(E, \mathcal{K}, P)$. Vom presupune în cele ce urmează că aceste variabile aleatoare au dispersii finite. Pentru simplitate vom utiliza notațiile

$$a_j = M(\tilde{x}_j), \quad \sigma_j^2 = V(\tilde{x}_j),$$

$$a_{(n)} = \sum_{j=1}^{n} a_j, \quad \sigma_{(n)}^2 = \sum_{j=1}^{n} \sigma_j^2 \tag{10.4.1}$$

și

$$\tilde{x}_{(n)} = (\sigma_{(n)}^2)^{-1} \sum_{j=1}^{n} (\tilde{x}_j - a_j). \tag{10.4.2}$$

Problema care se pune este următoarea: ce condiții trebuie impuse șirului de variabile aleatoare $(\tilde{x}_n)$ pentru ca șirul $(\tilde{x}_{(n)})$ definit de (10.4.2) să conveargă în repartiție spre o variabilă aleatoare normală? Această problemă, numită și *problema asimptotică centrală*, are o importanță deosebită în aplicațiile teoriei probabilităților permițînd asimilarea funcției de repartiție a lui $\tilde{x}_{(n)}$, pentru n suficient de mare, cu funcția de repartiție a unei variabile aleatoare normale.

Una dintre condițiile care s-a dovedit atît necesară cît și suficientă pentru ca $(\tilde{x}_{(n)})$ să conveargă în repartiție spre o variabilă aleatoare normală standard este condiția dată de Lindeberg, pe care o vom numi și *condiția* (L). Ea are următorul enunț: pentru orice $\varepsilon > 0$ avem

$$\lim_n (\sigma_{(n)}^2)^{-1} \sum_{j=1}^{n} \int_{|x-a_j|>\varepsilon\sigma_{(n)}} (x - a_j)^2 \, dF_j(x) = 0, \tag{10.4.3}$$

unde F_j este funcția de repartiție a variabilei aleatoare $\tilde{x}_j$.

Pentru a clarifica semnificația condiției lui Lindeberg, să fixăm ε și n și să considerăm evenimentele $A_j(\varepsilon)$, $1 \leqslant j \leqslant n$, definite astfel:

$$A_j(\varepsilon) = \{\omega : |\tilde{x}_j(\omega) - a_j| > \varepsilon\sigma_{(n)}\}. \tag{10.4.4}$$

Este clar că

$$P(A_j(\varepsilon)) = \int_{|x-a_j|>\varepsilon\sigma_{(n)}} \mathrm{d}F_j(x) \leqslant (\varepsilon^2\sigma_{(n)}^2)^{-1} \int_{|x-a_j|>\varepsilon\sigma_{(n)}} (x-a_j)^2 \,\mathrm{d}F_j(x) \quad (10.4.5)$$

Pe de altă parte, avem evident

$$P(\max_{1\leqslant j\leqslant n} |\widetilde{x}_j - a_j| > \varepsilon\sigma_{(n)}) = P\left(\bigcup_{j=1}^{n} A_j(\varepsilon)\right) \leqslant \sum_{j=1}^{n} P(A_j(\varepsilon)). \quad (10.4.6)$$

Din relațiile (10.4.5) și (10.4.6) se obține

$$P(\max_{1\leqslant j\leqslant n} |\widetilde{x}_j - a_j| > \varepsilon\sigma_{(n)}) \leqslant (\varepsilon^2\sigma_{(n)}^2) \sum_{j=1}^{n} \int_{|x-a_j|>\varepsilon\sigma_{(n)}} (x-a_j)^2 \,\mathrm{d}F_j(x)$$

și deci, conform condiției lui Lindeberg (10.4.3), rezultă că

$$\lim_n P(\max_{1\leqslant j\leqslant n} |\widetilde{x}_j - a_j| > \varepsilon\sigma_{(n)}) = 0 \quad (10.4.7)$$

pentru orice $\varepsilon > 0$. Această ultimă condiție arată prin urmare că termenii sumei (10.4.2) sînt „mici în mod uniform".

10.5. TEOREMA LIMITĂ CENTRALĂ

Vom prezenta mai întîi cîteva teoreme care dau condiții suficiente pentru ca șirul de variabile aleatoare $(\widetilde{x}_{(n)})$ definit de relația (10.4.2) să conveargă în repartiție spre o variabilă aleatoare normală standard. Cu probleme de acest tip s-au ocupat Lindeberg, Leapunov, Markov, lucrările lor fiind generate de rezultatul particular cunoscut sub numele de teorema lui DeMoivre-Laplace.

În demonstrația teoremei lui Lindeberg folosim următorul rezultat din analiza matematică.

Lema 10.5.1. Pentru orice $x \in R$ avem

$$\left| e^{ix} - \sum_{j=0}^{n} \frac{(ix)^j}{j!} \right| \leqslant \frac{|x|^{n+1}}{(n+1)!}.$$

Demonstrație. Pentru $n = 0$ avem

$$|e^{ix} - 1| = \left| \int_0^x e^{it}\, dt \right| \leqslant |x|.$$

Dacă presupunem rezultatul adevărat pentru n, atunci

$$\left| e^{ix} - \sum_{j=0}^{n+1} \frac{(ix)^j}{j!} \right| = \left| \int_0^x \left(e^{it} - \sum_{j=0}^{n} \frac{(it)^j}{j!} \right) dt \right| \leqslant$$

$$\leqslant \int_0^x \left| e^{it} - \sum_{j=0}^{n} \frac{(it)^j}{j!} \right| dt \leqslant \int_0^x \frac{|t|^{n+1}}{(n+1)!}\, dt = \frac{|x|^{n+2}}{(n+2)!},$$

ceea ce arată că inegalitatea este adevărată și pentru $n+1$ și demonstrația prin inducție a lemei este terminată.

Teorema 10.5.1 *(Lindeberg). Fie* $(\tilde{x}_n)$ *un șir de variabile aleatoare independente. Dacă este îndeplinită condiția lui Lindeberg* (10.4.3), *atunci șirul de variabile aleatoare* $(\tilde{x}_{(n)})$ *definit de* (10.4.2) *converge în repartiție către o variabilă aleatoare normală standard. În plus, avem*

$$\lim_n \max_{1 \leqslant j \leqslant n} [(\sigma_j^2)/(\sigma_{(n)}^2)] = 0 \qquad (10.5.1)$$

Demonstrație. Pentru orice j, $1 \leqslant j \leqslant n$, putem scrie

$$\sigma_j^2 = \int_{|x-a_j| \leqslant \varepsilon\sigma_{(n)}} (x - a_j)^2\, dF_j(x) + \int_{|x-a_j| > \varepsilon\sigma_{(n)}} (x - a_j)^2\, dF_j(x) \leqslant$$

$$\leqslant \varepsilon\sigma_{(n)}^2 + \sum_{j=1}^{n} \int_{|x-a_j| > \varepsilon\sigma_{(n)}} (x - a_j)^2\, dF_j(x),$$

de unde obținem imediat

$$\max_{1 \leqslant j \leqslant n} [(\sigma_j^2)/(\sigma_{(n)}^2)] \leqslant \varepsilon + (\sigma_{(n)}^2)^{-1} \sum_{j=1}^{n} \int_{|x-a_j| > \varepsilon\sigma_{(n)}} (x - a_j)^2\, dF_j(x). \qquad (10.5.2)$$

Scanned with OKEN Scanner

Trecînd la limită pentru $n \to +\infty$ în inegalitatea (10.5.2) și ținînd seama de (10.4.3), rezultă imediat relația căutată (10.5.1).

Fie acum variabilele aleatoare $\widetilde{x}_{nj}$ definite de relațiile

$$\widetilde{x}_{nj} = \frac{\widetilde{x}_j - a_j}{\sigma_{(n)}}, \quad 1 \leqslant j \leqslant n,$$

și fie F_{nj} funcțiile de repartiție corespunzătoare. Este clar că avem

$$M(\widetilde{x}_{nj}) = 0, \quad V(\widetilde{x}_{nj}) = (\sigma_j^2)/(\sigma_{(n)}^2). \tag{10.5.3}$$

Mai mult, prin schimbarea de variabilă $x - a_j = y\sigma_{(n)}$ integrala din (10.4.3) poate fi scrisă sub forma

$$\int_{|x-a_j|>\varepsilon\sigma_{(n)}} (x - a_j)^2 \, \mathrm{d}F_j(x) = \sigma_{(n)}^2 \int_{|y|>\varepsilon} y^2 \, \mathrm{d}F_{nj}(y).$$

Rezultă atunci că (10.4.3) poate fi pusă sub forma

$$\lim_n \sum_{j=1}^{n} \int_{|x|>\varepsilon} x^2 \, \mathrm{d}F_{nj}(x) = 0. \tag{10.5.4}$$

Deoarece avem evident

$$\widetilde{x}_{(n)} = \sum_{j=1}^{n} \widetilde{x}_{nj},$$

iar $\widetilde{x}_{n1}, \ldots, \widetilde{x}_{nn}$ sînt variabile aleatoare independente, rezultă că funcția caracteristică φ_n a variabilei aleatoare $\widetilde{x}_{(n)}$ poate fi scrisă în forma

$$\varphi_n(t) = \prod_{j=1}^{n} \varphi_{nj}(t), \tag{10.5.5}$$

unde φ_{nj} este funcția caracteristică a variabilei aleatoare $\widetilde{x}_{nj}$.

Pentru demonstrarea teoremei va fi suficient să arătăm, conform teoremei (8.5.3), că șirul de funcții caracteristice (φ_n) converge uniform, în orice interval mărginit, spre funcția caracteristică $e^{-t2/2}$ a unei variabile aleatoare normale standard.

Scanned with OKEN Scanner

Vom arăta mai întîi că φ_{nj} converge, cînd $n \to +\infty$, către 1 în orice interval mărginit, uniform în raport cu j. Într-adevăr, deoarece $M(\tilde{x}_{nj}) = 0$, rezultă că

$$\varphi_{nj}(t) - 1 = \int_R (e^{itx} - 1 - itx)\, dF_{nj}(x)$$

și deci, conform lemei (10.5.1), avem

$$|\varphi_{nj}(t) - 1| \leqslant (t^2/2)\int_R x^2\, dF_{nj}(x) = (t^2/2)\left[\int_{|x|\leqslant\varepsilon} x^2\, dF_{nj}(x) + \int_{|x|>\varepsilon} x^2\, dF_{nj}(x)\right] \leqslant$$

$$\leqslant (t^2/2)\left[\varepsilon^2 + \int_{|x|>\varepsilon} x^2\, dF_{nj}(x)\right]. \qquad (10.5.6)$$

Din (10.5.4) rezultă însă că pentru n suficient de mare putem scrie

$$\int_{|x|>\varepsilon} x^2\, dF_{nj}(x) < \varepsilon^2 \qquad (10.5.7)$$

pentru toți j, $1 \leqslant j \leqslant n$. Din (10.5.6) și (10.5.7) urmează atunci că, pentru n suficient de mare, putem scrie

$$|\varphi_{nj}(t) - 1| \leqslant \varepsilon^2 T^2, \qquad (10.5.8)$$

uniform în raport cu j, $1 \leqslant j \leqslant n$, și pentru orice interval finit $|t| \leqslant T$ cu $T > 0$. Din (10.5.8) rezultă atunci

$$\lim_n \varphi_{nj}(t) = 1,$$

uniform în raport cu j. Tot din (10.5.8) urmează că, pentru n suficient de mare,

$$|\varphi_{nj}(t) - 1| < 1/2 \qquad (10.5.9)$$

uniform în raport cu j, $1 \leqslant j \leqslant n$, în intervalul $|t| \leqslant T$.

Scanned with OKEN Scanner

Dezvoltînd în serie $\ln \varphi_n(t)$ și ținînd seama de (10.5.5) și (10.5.9), obținem *:

$$\ln \varphi_n(t) = \sum_{j=1}^{n} \ln \varphi_{nj}(t) = \sum_{j=1}^{n} \ln (1 + (\varphi_{nj}(t) - 1)) = \sum_{j=1}^{n} (\varphi_{nj}(t) - 1) + \rho_n, \tag{10.5.10}$$

unde

$$\rho_n = \sum_{j=1}^{n} \sum_{k=2}^{\infty} \frac{(-1)^{k-1}}{k} (\varphi_{nj}(t) - 1)^k.$$

Ținînd seama de (10.5.9), rezultă că avem evaluarea

$$|\rho_n| \leqslant \sum_{j=1}^{n} \sum_{k=2}^{\infty} (1/2) |\varphi_{nj}(t) - 1|^k = (1/2) \sum_{j=1}^{n} \frac{|\varphi_{nj}(t) - 1|^2}{1 - |\varphi_{nj}(t) - 1|} \leqslant$$

$$\leqslant \sum_{j=1}^{n} |\varphi_{nj}(t) - 1|^2 \leqslant \max_{1 \leqslant j \leqslant n} |\varphi_{nj}(t) - 1| \sum_{j=1}^{n} |\varphi_{nj}(t) - 1|.$$

Deoarece

$$\sum_{j=1}^{n} |\varphi_{nj}(t) - 1| \leqslant \sum_{j=1}^{n} \int_R |e^{itx} - 1 - itx| \quad dF_{nj}(x) \leqslant$$

$$\leqslant (t^2/2) \sum_{j=1}^{n} \int_R x^2 \, dF_{nj}(x) = (t^2/2) \sum_{j=1}^{n} V(\tilde{x}_{nj}) = t^2/2$$

rezultă că

$$|\rho_n| \leqslant (t^2/2) \max_{1 \leqslant j \leqslant n} |\varphi_{nj}(t) - 1| \tag{10.5.11}$$

și deci

$$\lim_n \rho_n = 0 \tag{10.5.12}$$

uniform în orice interval $|t| \leqslant T$.

---

* Am notat cu $\ln \varphi_n(t)$ *valoarea principală* a logaritmului.

Pe de altă parte, punînd

$$\sum_{j=1}^{n} (\varphi_{nj}(t) - 1) = -(t^2/2) + \rho_n', \tag{10.5.13}$$

unde

$$\rho_n' = t^2/2 + \sum_{j=1}^{n} (\varphi_{nj}(t) - 1) = t^2/2 + \sum_{j=1}^{n} \int_R (e^{itx} - 1 - itx)\, dF_{nj}(x) =$$

$$= \sum_{j=1}^{n} \int_{|x| \leqslant \varepsilon} \left(e^{itx} - 1 - itx - \frac{(itx)^2}{2} \right) dF_{nj}(x) +$$

$$+ \sum_{j=1}^{n} \int_{|x| > \varepsilon} \left(e^{itx} - 1 - itx - \frac{(itx)^2}{2} \right) dF_{nj}(x)$$

și aplicînd din nou lema (10.5.1), rezultă că

$$|\rho_n'| \leqslant (|t|^3/6) \sum_{j=1}^{n} \int_{|x| \leqslant \varepsilon} |x|^3\, dF_{nj}(x) + t^2 \sum_{j=1}^{n} \int_{|x| > \varepsilon} x^2\, dF_{nj}(x) \leqslant$$

$$\leqslant (|t|^3 \varepsilon/6) \sum_{j=1}^{n} \int_{|x| \leqslant \varepsilon} x^2\, dF_{nj}(x) + t^2 \sum_{j=1}^{n} \int_{|x| > \varepsilon} x^2\, dF_{nj}(x) =$$

$$= (|t|^3/6) + t^2(1 - (|t|\varepsilon)/6) \sum_{j=1}^{n} \int_{|x| > \varepsilon} x^2\, dF_{nj}(x). \tag{10.5.14}$$

Din (10.5.14), ținînd seama de condiția lui Lindeberg, se deduce ușor că

$$\lim_n \rho_n' = 0, \tag{10.5.15}$$

uniform în orice interval $|t| \leqslant T$.

Din (10.5.10) și (10.5.13) rezultă că putem scrie

$$\ln \varphi_n(t) = -t^2/2 + \rho_n + \rho_n'. \tag{10.5.16}$$

Scanned with OKEN Scanner

Trecînd la limită în (10.5.16) pentru $n \to +\infty$ și ținînd seama de (10.5.12) și (10.5.15), obținem

$$\lim_n \ln \varphi_n(t) = -t^2/2$$

sau

$$\lim_n \varphi_n(t) = e^{-t2/2},$$

uniform în orice interval $|t| \leqslant T$, ceea ce demonstrează complet teorema.

Din teorema lui Lindeberg obținem ușor următoarele consecințe.

Teorema 10.5.2 *(Leapunov). Fie* $(\widetilde{x}_n)$ *un șir de variabile aleatoare independente. Dacă există un număr* $\delta > 0$ *astfel încît*

$$\lim_n (\sigma_{(n)})^{-2-\delta} \sum_{j=1}^{n} M(|\widetilde{x}_j - a_j|^{2+\delta}) = 0, \qquad (10.5.17)$$

atunci șirul $(\widetilde{x}_{(n)})$ *definit de* (10.5.2) *converge în repartiție spre o variabilă aleatoare normală standard.*

Demonstrație. Vom arăta că (10.5.17), care mai este numită și *condiția lui Leapunov*, implică condiția lui Lindeberg (10.4.3). Într-adevăr, pentru orice $\varepsilon > 0$ și orice n, avem

$$(\sigma^2_{(n)})^{-1} \sum_{j=1}^{n} \int_{|x-a_j|>\varepsilon\sigma(n)} (x-a_j)^2 \, dF_j(x) \leqslant$$

$$\leqslant (\sigma_{(n)})^{-2-\delta} \, \varepsilon^{-\delta} \sum_{j=1}^{n} \int_{|x-a_j|>\varepsilon\sigma(n)} |x-a_j|^{2+\delta} \, dF_j(x) \leqslant$$

$$\leqslant (\sigma_{(n)})^{-2-\delta} \, \varepsilon^{-\delta} \sum_{j=1}^{n} M(|\widetilde{x}_j - a_j|^{2+\delta}),$$

de unde se deduce imediat că (10.4.3) este îndeplinită.

Alte condiții suficiente utile în aplicații sînt date în teorema următoare.

Teorema 10.5.3 *Fie* $(\widetilde{x}_n)$ *un șir de variabile aleatoare independente. Fiecare dintre condițiile următoare este suficientă pentru ca șirul de variabile aleatoare* $(\widetilde{x}_{(n)})$ *să conveargă în repartiție spre o variabilă aleatoare normală standard:*

(*a*) *Există momentele centrate absolute de ordinul trei*

$$\rho_j^3 = M(|\widetilde{x}_j - a_j|^3)$$

și

$$\lim_n (\rho_{(n)})/(\sigma_{(n)}) = 0, \tag{10.5.18}$$

unde

$$\rho_{(n)}^3 = \sum_{j=1}^{n} \rho_j^3;$$

(*b*) *Variabilele aleatoare* $(\widetilde{x}_n)$ *admit dispersii finite, sînt uniform mărginite și*

$$\lim_n \sigma_{(n)} = +\infty; \tag{10.5.19}$$

(*c*) *Variabilele aleatoare* $(\widetilde{x}_n)$ *sînt identic repartizate și admit dispersii finite.*

Demonstrație. (*a*) Condiția (10.5.18) implică evident condiția lui Leapunov pentru $\delta = 1$.

(*b*) Deoarece $(\widetilde{x}_j)$ sînt uniform mărginite, rezultă că există $M > 0$ astfel încît

$$|\widetilde{x}_j - a_j| \leqslant M \tag{10.5.20}$$

pentru orice j natural. Putem deci scrie

$$\int_{|x-a_j| \geqslant \varepsilon\sigma_{(n)}} (x-a_j)^2 \, dF_j(x) \leqslant M^2 \int_{|x-a_j| \geqslant \varepsilon\sigma_{(n)}} dF_j(x) =$$

$$= M^2 P(|\widetilde{x}_j - a_j| \geqslant \varepsilon\sigma_{(n)}). \tag{10.5.21}$$

Ținînd seama de condiția (10.5.19), rezultă că pentru n suficien de mare, avem $\varepsilon\sigma_{(n)} > M$ și deci (10.5.20) implică

$$P(|\widetilde{x}_j - a_j| \geqslant \varepsilon\sigma_{(n)}) = 0. \tag{10.5.22}$$

Scanned with OKEN Scanner

Din (10.5.21) și (10.5.22) rezultă

$$\int\limits_{|x-a_j|\geqslant\varepsilon\sigma_{(n)}} (x-a_j)^2\,\mathrm{d}F_j(x)=0, \qquad 1\leqslant j\leqslant n,$$

pentru n suficient de mare, ceea ce implică evident condiția lui Lindeberg (10.4.3).

(*c*) Deoarece $(\widetilde{x}_j)$ sînt identic repartizate, avem $F_j=F$, $a_j=a$ și $\sigma_j^2=\sigma^2$ pentru toți j. Rezultă atunci $\sigma_{(n)}^2=n\sigma^2$ și deci

$$(\sigma_{(n)}^2)^{-1}\sum_{j=1}^{n}\int\limits_{|x-a_j|>\varepsilon\sigma_{(n)}}(x-a_j)^2\,\mathrm{d}F_j(x)=(1/n\sigma^2)\sum_{j=1}^{n}\int\limits_{|x-a|>\varepsilon\sigma\sqrt{n}}(x-a)^2\,\mathrm{d}F(x)=$$

$$=(1/\sigma^2)\int\limits_{|x-a|>\varepsilon\sigma\sqrt{n}}(x-a)^2\,\mathrm{d}F(x). \qquad (10.5.23)$$

Din faptul că $(\widetilde{x}_j)$ admit dispersii finite, rezultă că ultima integrală obținută în (10.5.23) tinde spre zero cînd $n\to+\infty$, ceea ce înseamnă, conform cu (10.5.23), că este satisfăcută condiția lui Lindeberg.

Vom arăta în cele ce urmează că condiția lui Lindeberg nu este numai suficientă, dar și necesară pentru ca șirul de variabile aleatoare $(\widetilde{x}_{(n)})$ să conveargă în repartiție spre o variabilă normală standard.

Teorema 10.5.4 *(Feller). Fie* $(\widetilde{\mathrm{x}}_{\mathrm{n}})$ *un șir de variabile aleatoare independente. Dacă șirul* $(\widetilde{\mathrm{x}}_{(\mathrm{n})})$ *dat de* (10.4.2) *converge în repartiție spre o variabilă aleatoare normală standard și dacă este îndeplinită condiția* (10.5.1), *atunci este îndeplinită condiția lui Lindeberg* (10.4.3).

Demonstrație. Vom relua notațiile din teorema 10.5.1. Deoarece

$$\varphi_{nj}(t)=\int_R(\mathrm{e}^{itx}-itx)\,\mathrm{d}F_{nj}(x),$$

după o dezvoltare în serie a lui e^{itx}, obținem

$$\varphi_{nj}(t)=1-\theta_j(t^2/2)(\sigma_j^2/\sigma_{(n)}^2), \quad 1\leqslant j\leqslant n.$$

Din condiția (10.5.1) rezultă atunci că

$$\lim_n \max_{1\leqslant j\leqslant n} |\varphi_{nj}(t) - 1| = 0 \tag{10.5.24}$$

și

$$\lim_n \sum_{j=1}^{n} |\varphi_{nj}(t) - 1| = 0. \tag{10.5.25}$$

Prin ipoteză,

$$\lim_n \varphi_n(t) = \lim_n \prod_{j=1}^{n} \varphi_{nj}(t) = e^{-t^2/2},$$

de unde

$$\lim_n \sum_{j=1}^{n} \ln \varphi_{nj}(t) = -t^2/2, \tag{10.5.26}$$

dacă se ține seama și de faptul că, pentru n suficient de mare, $\ln \varphi_{nj}(t)$ există.

Deoarece

$$\ln z = (z-1) + \theta |z-1|^2,$$

rezultă că, pentru n suficient de mare, putem scrie

$$\left|(t^2/2) - \sum_{j=1}^{n} (1 - \varphi_{nj}(t))\right| \leqslant \left|(t^2/2) + \sum_{j=1}^{n} \ln \varphi_{nj}(t)\right| + \theta \sum_{j=1}^{n} |\varphi_{nj}(t) - 1|^2 \leqslant$$

$$\leqslant \left|(t^2/2) + \sum_{j=1}^{n} \ln \varphi_{nj}(t)\right| + \theta \max_{1\leqslant j\leqslant n} \sum_{j=1}^{n} |\varphi_{nj}(t) - 1|. \tag{10.5.27}$$

Trecînd la limită în (10.5.27) pentru $n \to +\infty$ și ținînd seama de (10.5.24), (10.5.25) și (10.5.26), rezultă imediat că

$$\lim_n |\psi_n(t)| = 0, \tag{10.5.28}$$

unde

$$\psi_n(t) = (t^2/2) - \sum_{j=1}^{n} (1 - \varphi_{nj}(t)) = (t^2/2) - \sum_{j=1}^{n} \int_R (1 - e^{itx})\, dF_{nj}(x). \tag{10.5.29}$$

Să observăm însă că putem scrie

$$\psi_n(t) = (t^2/2) - \sum_{j=1}^{n} \int_R (1 - \cos tx)\, dF_{nj}(x) + i \sum_{j=1}^{n} \int_R \sin tx\, dF_{nj}(x)$$

și deci partea reală a lui $\psi_n(t)$ se poate pune, pentru orice $\varepsilon > 0$, sub forma

$$\operatorname{Re}(\psi_n(t)) = (t^2/2) - \sum_{j=1}^{n} \int_{|x| \leqslant \varepsilon} (1 - \cos tx)\, dF_{nj}(x) -$$

$$- \sum_{j=1}^{n} \int_{|x| > \varepsilon} (1 - \cos tx)\, dF_{nj}(x). \qquad (10.5.30)$$

Dar

$$\int_{|x| \leqslant \varepsilon} (1 - \cos tx)\, dF_{nj}(x) = 2 \int_{|x| \leqslant \varepsilon} \sin^2(tx/2)\, dF_{nj}(x) \leqslant$$

$$\leqslant (t^2/2) \int_{|x| \leqslant \varepsilon} x^2\, dF_{nj}(x) = (t^2/2)\left[V(\tilde{x}_{nj}) - \int_{|x| > \varepsilon} x^2\, dF_{nj}(x) \right]$$

și deci

$$\sum_{j=1}^{n} \int_{|x| \leqslant \varepsilon} (1 - \cos tx)\, dF_{nj}(x) \leqslant (t^2/2)\left[1 - \sum_{j=1}^{n} \int_{|x| > \varepsilon} x^2\, dF_{nj}(x) \right]. \qquad (10.5.31)$$

Pe de altă parte, avem

$$\sum_{j=1}^{n} \int_{|x| > \varepsilon} (1 - \cos tx)\, dF_{nj}(x) \leqslant 2 \sum_{j=1}^{n} \int_{|x| > \varepsilon} dF_{nj}(x) \leqslant$$

$$\leqslant 2\varepsilon^{-2} \sum_{j=1}^{n} \int_{|x| > \varepsilon} x^2\, dF_{nj}(x) \leqslant 2\varepsilon^{-2} \sum_{j=1}^{n} V(\tilde{x}_{nj}) = 2/\varepsilon^2. \qquad (10.5.32)$$

Scanned with OKEN Scanner

Din relațiile (10.5.30), (10.5.31) și (10.5.32) se deduce ușor

$$|\mathrm{Re}\,(\psi_n(t))| \geqslant t^2/2 - t^2/2 + (t^2/2)\sum_{j=1}^{n}\int_{|x|>\varepsilon} x^2\,\mathrm{d}F_{nj}(x) - 2/\varepsilon^2$$

sau

$$0 \leqslant \sum_{j=1}^{n}\int_{|x|>\varepsilon} x^2\,\mathrm{d}F_{nj}(x) \leqslant (2/t^2)\,[|\mathrm{Re}\,(\psi_n(t))| + 2/\varepsilon^2]. \qquad (10.5.33)$$

Trecînd la limită în (10.5.33) mai întîi pentru $n \to +\infty$ și apoi pentru $t \to +\infty$ și ținînd seama de (10.5.30), rezultă

$$\lim_n \sum_{j=1}^{n}\int_{|x|>\varepsilon} x^2\,\mathrm{d}F_{nj}(x) = 0$$

și deci condiția lui Lindeberg este verificată (vezi forma (10.5.4) a acestei condiții).

Teoremele de mai sus se pot extinde ușor la șiruri de vectori aleatori.

10.6. TEOREME LIMITĂ LOCALE

Teoremele din paragraful precedent stabilesc condițiile în care un șir de funcții de repartiție $(F_{(n)})$ converge în repartiție spre funcția de repartiție $F_N(x|0,1)$. Dacă funcțiile de repartiție au densități de probabilitate sau funcții de masă, atunci se pune problema dacă șirul densităților de probabilitate sau al funcțiilor de masă converge spre densitatea de probabilitate normală standard $f_N(x|0,1)$. Teoremele care stabilesc condițiile în care această convergență are loc sînt numite *teoreme limită locale*, spre a le deosebi de teoremele limită *(globale)* privind funcțiile de repartiție, care au fost prezentate în § 10.5.

Vom considera mai întîi variabile aleatoare discrete, care au repartiții de un tip special.

Definiția 10.6.1. *O variabilă aleatoare discretă* $\tilde{x}$ *are o repartiție în rețea, dacă există un număr real* a *și un număr pozitiv* h *cu proprietatea că toate valorile posibile ale lui* $\tilde{x}$ *sînt de forma* $a + kh$ *cu* $k \in Z$.

Numărul $h > 0$ este numit *pasul rețelei.*

O caracterizare a repartițiilor în rețea este dată în teorema care urmează.

Teorema 10.6.1. *Condiția necesară și suficientă ca variabila aleatoare* $\tilde{x}$ *să aibă o repartiție în rețea este ca să existe o valoare* $t \neq 0$ *pentru care funcția caracteristică a lui* $\tilde{x}$ *să fie în modul egală cu* 1.

Demonstrație. Să arătăm mai întîi că această condiție este necesară. Dacă $\tilde{x}$ are repartiție în rețea, atunci există $a \in R$ și $h > 0$ astfel încît toate valorile posibile ale lui $\tilde{x}$ să fie de forma $a + kh$, $k \in Z$. Funcția caracteristică a lui $\tilde{x}$ este atunci

$$\varphi_{\tilde{x}}(t) = \sum_{k \in Z} e^{it(a+kh)} f_{\tilde{x}}(a + kh) = e^{iat} \sum_{k \in Z} e^{ithk} f_{\tilde{x}}(a + kh)$$

și deci

$$\varphi_{\tilde{x}}(2\pi/h) = e^{ia(2\pi/h)} \sum_{k \in Z} f_x(a + kh) = e^{ia(2\pi/h)},$$

de unde, evident,

$$|\varphi_{\tilde{x}}(2\pi/h)| = 1, \tag{10.6.1}$$

ceea ce arată că valoarea căutată a lui t este $t = 2\pi/h$.

Pentru demonstrarea suficienței, fie $t_0 \neq 0$ pentru care $|\varphi(t_0)| = 1$. Există atunci $\alpha \in [0, 2\pi]$ astfel încît $\varphi(t_0) = e^{i\alpha}$ și deci

$$\int_R e^{it_0 x} \, dF(x) = e^{i\alpha},$$

unde F este funcția de repartiție a lui $\tilde{x}$. Rezultă atunci

$$\int_R e^{i(t_0 x - \alpha)} \, dF(x) = 1$$

și deci, luînd partea reală, avem

$$\int_R \cos(t_0 x - \alpha) \, dF(x) = 1.$$

Scanned with OKEN Scanner

Această relație nu poate avea loc decît dacă F are puncte de creștere numai în acele puncte $x \in R$ pentru care

$$\cos(t_0 x - \alpha) = 1,$$

adică $x \in R$ sînt de forma

$$x = \alpha/t_0 + 2\pi k/t_0, \quad k \in Z. \tag{10.6.2}$$

Rezultă deci că valorile posibile ale variabilei aleatoare $\tilde{x}$ sînt de forma $a+kh$ cu $a = \alpha/t_0$ și $h = |2\pi/t_0|$, adică tocmai ceea ce trebuia demonstrat.

Definiția 10.6.2. *Pasul* $h > 0$ *al unei repartiții în rețea este maximal dacă nu există* $b \in R$ *și* $h' > h$ *astfel încît valorile posibile ale variabilei aleatoare să fie de forma* $b+kh'$ *cu* $k \in Z$.

Condiții echivalente ca pasul unei repartiții în rețea să fie maximal sînt date în teorema următoare.

Teorema 10.6.2. *Fiecare dintre condițiile următoare sînt necesare și suficiente ca pasul* h *al unei repartiții în rețea să fie maximal:*

(*a*) *Cel mai mare divizor comun al tuturor perechilor de diferențe dintre două valori posibile ale lui* $\tilde{x}$ *divizate prin* h *să fie egal cu* 1;

(*b*) *Pentru orice* t *pentru care* $0 < |t| < (2\pi)/h$ *avem* $|\varphi_{\tilde{x}}(t)| < 1$, *iar* $|\varphi_{\tilde{x}}(2\pi/h)| = 1$.

Demonstrație. (*a*) Această afirmație este imediată.

(*b*) Să arătăm mai întîi necesitatea condiției. Dacă ar exista t_0 cu $0 < |t_0| < 2\pi/h$ pentru care $|\varphi(t_0)| = 1$, atunci conform demonstrației teoremei precedente, rezultă că $h' = |2\pi/t_0|$ este pasul repartiției, ceea ce arată că avem $h' > h$ și deci h nu ar fi maximal.

Suficiența condiției rezultă imediat: dacă h nu este maximal, atunci există $h' > h$ care poate fi luat drept pas al repartiției și deci, conform cu (10.6.1), rezultă că $|\varphi_{\tilde{x}}(2\pi/h')| = 1$, care, împreună cu $2\pi/h' < 2\pi/h$, contrazice condiția (*b*).

Observație. Din teorema (10.6.2) rezultă că, dacă h este un pas maximal, atunci pentru orice $\varepsilon > 0$ există un număr $\alpha = \alpha(\varepsilon)$ cu proprietatea că

$$|\varphi(t)| \leqslant e^{-\alpha} \tag{10.6.3}$$

pentru orice t cu proprietatea $\varepsilon \leqslant |t| \leqslant (2\pi/h) - \varepsilon$.

Scanned with OKEN Scanner

Să revenim la problema pe care ne-am pus-o inițial. Să considerăm un șir de variabile aleatoare independente și identic repartizate $(\tilde{x}_n)$, avînd repartiție în rețea cu valorile posibile de forma $a + kh$, $k \in Z$. Dacă $(\tilde{y}_n)$ este șirul de variabile aleatoare definite de

$$\tilde{y}_n = \tilde{x}_1 + \ldots + \tilde{x}_n, \tag{10.6.4}$$

atunci este clar că $(\tilde{y}_n)$ sînt variabile aleatoare discrete care au de asemenea o repartiție în rețea, valorile posibile ale variabilei $\tilde{y}_n$ fiind de forma $na + kh$, $k \in Z$.

Vom nota

$$M(\tilde{x}_n) = m, \quad V(\tilde{x}_n) = \sigma^2 \tag{10.6.5}$$

și

$$a_{(n)} = M(\tilde{y}_n) = nm; \quad \sigma^2_{(n)} = V(\tilde{y}_n) = n\sigma^2. \tag{10.6.6}$$

În sfîrșit, vom nota

$$z_{nk} = \frac{na + kh - a_{(n)}}{\sigma_{(n)}} = \frac{na + kh - nm}{\sigma\sqrt{n}}. \tag{10.6.7}$$

Teorema 10.6.3 *(Gnedenko). Fie* $(\tilde{x}_n)$ *un șir de variabile aleatoare independente și identic repartizate, avînd repartiția în rețea cu valorile posibile de forma* $a + kh$, $k \in Z$, *iar media* $m = M(\tilde{x}_n)$ *și dispersia* $\sigma^2 = V(\tilde{x}_n)$ *sînt finite.*

Condiția necesară și suficientă ca

$$\lim_n \left[(\sigma_{(n)}/h) f_{\tilde{y}_n}(na + kh) - (2\pi)^{-1/2} e^{-z^2_{nk}/2}\right] = 0, \tag{10.6.8}$$

uniform în raport cu $k \in Z$, *este ca pasul* h *să fie maximal.*

Demonstrație. Nu prezentăm aici demonstrația acestei teoreme, care poate fi găsită în Gnedenko, B. (1969).

Scanned with OKEN Scanner

Celălalt caz interesant, pe care l-am menționat deja, este cel al șirului de funcții de repartiție care are densități de probabilitate. Prezentăm, fără demonstrație, un rezultat în acest sens.

Teorema 10.6.4. *Fie* $(\widetilde{x}_n)$ *un șir de variabile aleatoare independente și identic repartizate cu media și dispersia finite.*

Să presupunem, de asemenea, că începînd de la un anumit număr natural n_0, *variabilele aleatoare* $(\widetilde{x}_{(n)})$ *definite de relația* (10.4.2) *au densități de probabilitate* $f_{(n)}(x)$.

Condiția necesară și suficientă ca

$$\lim_n (f_{(n)}(x) - (2\pi)^{-1/2}e^{-x^2/2}) = 0$$

să aibă loc uniform în raport cu $x \in R$ *este ca să existe un număr natural* n_1 *astfel încît funcția* $f_{(n_1)}(x)$ *să fie mărginită.*

Capitolul 11

INTRODUCERE ÎN STATISTICA MATEMATICĂ

11.1. INTRODUCERE

Sîntem în acest moment suficient de bine pregătiți pentru a face un pas important înainte în studierea problemelor de decizie amintite în prima parte a lucrării. Este vorba de utilizarea informației disponibile pentru scopurile luării deciziilor.

Elementul principal cu care se lucrează în statistica matematică este o mulțime formată din obiecte sau indivizi (unități) care se diferențiază prin diverse *atribute* sau *caracteristici*, mulțime care este numită *colectivitate* sau *populație*. Pentru caracterizarea acestor mulțimi, statistica descriptivă clasică a utilizat diferite valori tipice asociate populațiilor, cum ar fi media, mediana, modul, momente de diverse ordine, cuantile etc. Aceste noțiuni au fost introduse în statistica matematică independent de teoria probabilităților.

Trecerea de la statistica descriptivă la studierea științifică a populațiilor statistice a fost marcată, la începutul secolului nostru, de K. Pearson și elevii săi, care au început să considere caracteristicele unei populații ca reprezentînd variabile aleatoare. Această nouă viziune a deschis drumul utilizării intense a teoriei probabilităților în problemele statisticii, dînd naștere la ceea ce se numește astăzi *statistica matematică*.

11.2. SELECȚIA ALEATOARE DINTR-O POPULAȚIE

Pentru fixarea ideilor, să considerăm că sîntem în prezența unei populații E finite și vrem să obținem informații despre anumite caracteristici ale indivizilor $\omega \in E$. Presupunem că este posibil să măsurăm k caracteristici $x_1(\omega), \ldots, x_k(\omega)$ pentru

Scanned with OKEN Scanner

fiecare individ $\omega \in E$. În sfîrșit, presupunem că putem efectua un *experiment* format din *observații* sau *probe* succesive, fiecare observație constînd în selectarea *în mod aleator* (la întîmplare) a unuia din indivizii eligibili la momentul respectiv și măsurarea celor k caracteristici. Cu alte cuvinte, fiecare *observație* sau *probă* de un anumit rang i constă din obținerea unui vector cu k componente $\mathbf{x}_i = (x_{1i}, \ldots, x_{ki})'$ format cu valorile numerice $x_{1i}, \ldots, x_{ki}$ ale caracteristicilor cercetate la această observație.

Evident, înainte de efectuarea observației de rang i, vectorul $\mathbf{x}_i$ este necunoscut și poate fi deci privit ca un vector aleator, pe care îl vom nota prin $\tilde{\mathbf{x}}_i$. Repartiția acestui vector aleator este determinată de modul în care se selectează indivizii pentru observare.

Există două metode principale de selecție, larg utilizate în practică. Prima metodă este numită *selecție aleatoare cu înlocuire* sau *selecție repetată:* la fiecare observație i, fiecare individ $\omega \in E$ are aceeași șansă de a fi selectat ca orice alt individ $\omega' \in E$, indiferent dacă ω sau (și) ω' au mai fost selectați anterior. Selecția aleatoare cu înlocuire conduce la vectori aleatori $\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \ldots$ care sînt evident independenți și identic repartizați, avînd funcția de masă*

$$f(\mathbf{x}) = P(\tilde{\mathbf{x}} = \mathbf{x}) = \frac{\operatorname{card}\{\omega : \tilde{\mathbf{x}}_i(\omega) = \mathbf{x}\}}{\operatorname{card} E} \qquad (11.2.1)$$

care este numită și *funcția de masă a populației***.

O altă metodă de selecție este *selecția aleatoare fără înlocuire* sau *selecția nerepetată:* la observația de rang i, individul $\omega \in E$ este eligibil pentru selectare dacă și numai dacă nu a fost selectat anterior, fiecare individ eligibil la etapa (momentul, pasul) i avînd aceeași șansă de a fi selectat ca orice alt individ *eligibil* la acest moment. Prin urmare la observația de rang i, numărul indivizilor eligibili este $\operatorname{card} E - (i - 1)$ și deci probabilitatea de a selecta un individ la etapa i este

$$\frac{\operatorname{card} E - (i - 1)}{\operatorname{card} E}. \qquad (11.2.2)$$

---

* Pentru orice mulțime A notăm prin $\operatorname{card} A$ numărul (finit sau transfinit) al elementelor lui A.

** Mulțimea E este presupusă finită, așa cum am precizat mai sus.

Scanned with OKEN Scanner

În acest caz, este clar că probabilitatea de a selecta un individ $\omega \in E$ la etapa i depinde atît de rangul observației, cît și de rezultatele obținute la observațiile anterioare. Rezultă atunci că vectorii aleatori $\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \ldots$ nu mai sînt independenți. Din acest motiv, vom considera în general experimente care conduc la vectori aleatori $\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \ldots$ dependenți*, cazul selecției repetate obținîndu-se imediat ca un caz particular.

Există cazuri în care mulțimea E de indivizi (sau de evenimente elementare, cum le-am numit în teoria probabilităților) nu este ușor de determinat. În cele mai multe cazuri, această specificare nu este însă necesară, experimentul putînd fi descris în mare parte prin șirul (finit sau infinit) al funcțiilor de masă condiționate

$$f(\mathbf{x}_1 \mid \boldsymbol{\theta}),\ f(\mathbf{x}_2 \mid \mathbf{x}_1,\ \boldsymbol{\theta}),\ \ldots,\ f(\mathbf{x}_i \mid \mathbf{x}_1,\ \ldots,\ \mathbf{x}_{i-1},\ \boldsymbol{\theta}), \qquad (11.2.3)$$

unde $\boldsymbol{\theta}$ este un vector conținînd drept componente mărimi asupra cărora urmează să se obțină informații în urma experimentului. Vectorul $\boldsymbol{\theta}$ va fi numit și *vector de stare*, conform terminologiei adoptată în prima parte a lucrării pentru elementele unei probleme de decizie.

Pentru a defini complet un experiment mai trebuie specificat și numărul n al observațiilor care trebuie efectuate și care mai este numit și *volumul selecției*. În multe experimente, volumul selecției este predeterminat. Există însă cazuri cînd numărul observațiilor nu este determinat dinainte cu certitudine. În astfel de cazuri, vom avea nevoie să cunoaștem probabilitatea de a face observația de rang $i + 1$, condiționată de rezultatele obținute la primele i observații. Aceste probabilități depind de vectorul de stare $\boldsymbol{\theta}$, dar, în general, pot depinde și de un alt vector pe care îl vom nota prin $\boldsymbol{\theta}_\alpha$: acesta conține drept componente mărimi de care depinde procesul de oprire al experimentului și care ar putea influența informațiile pe care le obținem despre vectorul de stare $\boldsymbol{\theta}$.

În cazul cînd E nu este o mulțime finită, putem relua construcția de mai sus, presupunînd că vectorii $\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \ldots$ au densități de probabilitate condiționate pe care le vom nota de asemenea prin $f(\mathbf{x}_1|\boldsymbol{\theta}), f(\mathbf{x}_2|\mathbf{x}_1, \boldsymbol{\theta}), \ldots, f(\mathbf{x}_i|\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, \boldsymbol{\theta}), \ldots$

---

* Această dependență este un exemplu de *dependență cu legături complete*, introdusă și studiată de școala românească de teoria probabilităților.

Scanned with OKEN Scanner

11.3. PROCESE GENERATOARE DE DATE ȘI PROCESE DE OPRIRE

Prin *proces generator de date* vom înțelege un mod de a construi un șir (finit sau infinit) de vectori aleatori $\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \ldots$ care să reprezinte rezultatele posibile ale unui experiment. Un proces generator de date este complet specificat de șirul $f(\mathbf{x}_1|\boldsymbol{\theta})$ $f(\mathbf{x}_2|\mathbf{x}_1, \boldsymbol{\theta}), \ldots$ al repartițiilor elementare* condiționate. Vom interpreta $f(\mathbf{x}_i|\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, \boldsymbol{\theta})$ drept repartiția elementară a lui $\tilde{\mathbf{x}}_i$ condiționată de $\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, \boldsymbol{\theta}$ și de faptul că observația de rang i este efectuată.

Prin *proces de oprire* înțelegem un șir (finit sau infinit) de *probabilități de oprire* $\pi(0 \mid \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$, $\pi(1 \mid \mathbf{x}_1, \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha), \ldots$ $\pi(i \mid \mathbf{x}_1, \ldots, \mathbf{x}_i, \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$, unde prin $\pi(i \mid \mathbf{x}_1, \ldots, \mathbf{x}_i, \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$ notăm probabilitatea ca experimentul să se termine după observația de rang i, condiționată de faptul că au fost efectuate i observații, au fost obținute valorile $\mathbf{x}_1, \ldots, \mathbf{x}_i$, vectorul de stare este $\boldsymbol{\theta}$, iar vectorul parametrilor adiționali este $\boldsymbol{\theta}_\alpha$. Este clar că $\pi(0|\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$ reprezintă probabilitatea ca să nu se facă nici o observație.

Este evident că probabilitatea $s(i \mid \mathbf{x}_1, \ldots, \mathbf{x}_i, \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$ ca experimentul să fie constituit din exact i observații este dată de relațiile

$$s(0|\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha) = \pi(0 \mid \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha), \tag{11.3.1}$$

$$s(1|\mathbf{x}_1, \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha) = \pi(1|\mathbf{x}_1, \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)\,(1 - \pi(0|\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)), \tag{11.3.2}$$

$$s(i|\mathbf{x}_1, \ldots, \mathbf{x}_i, \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha) =$$

$$= \pi(i|\mathbf{x}_1, \ldots, \mathbf{x}_i, \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha) \prod_{j=0}^{i-1} (1 - \pi(j|\mathbf{x}_1, \ldots, \mathbf{x}_j, \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)) \tag{11.3.3}$$

unde

$$\pi(0|\mathbf{x}_1, \mathbf{x}_0, \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha) = \pi(0|\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha).$$

Este clar că un *experiment statistic* este complet determinat de un proces generator de date și de un proces de oprire. Mai exact, dacă se consideră un număr aleator $\tilde{n}$ de vectori aleatori

---

* Acestea pot fi funcții de masă, densități de probabilitate sau repartiții elementare autentice (vezi § 4.3).

$\tilde{x}_1, \ldots, \tilde{x}_{\tilde{n}}$, atunci repartiția elementară a vectorului aleator $(\tilde{n}, \tilde{x}_1, \ldots, \tilde{x}_{\tilde{n}})$ condiționată de θ și θ_α se poate exprima în modul următor:

$$f(n, x_1, \ldots, x_n|\theta, \theta_\alpha) = \\ = s(n | x_1, \ldots, x_n, \theta, \theta_\alpha) f(x_1, \ldots, x_n|\theta), \tag{11.3.4}$$

unde

$$f(x_1, \ldots, x_n|\theta) = f(x_1|\theta) \prod_{i=2}^{n} f(x_i|x_1, \ldots, x_{i-1}, \theta) \tag{11.3.5}$$

este densitatea de probabilitate a vectorului aleator $(\tilde{x}_1, \ldots, \tilde{x}_n)$, condiționată de θ.

Formula (11.3.4) se obține în modul următor. Deoarece vectorul θ conține parametrii adiționali care sînt relevanți numai pentru procesul de oprire, rezultă că putem scrie

$$f(x_1, \ldots, x_n|\theta, \theta_\alpha) = f(x_1, \ldots, x_n|\theta)$$

pentru toți θ_α. Rezultă atunci că

$$f(n, x_1, \ldots, x_n|\theta, \theta_\alpha) = f(n|x_1, \ldots, x_n, \theta, \theta_\alpha) f(x_1, \ldots, x_n|\theta, \theta_\alpha),$$

de unde se obține imediat relația căutată (11.3.4).

Este clar de asemenea că în urma unui experiment statistic se obține un rezultat de forma $(\tilde{n}, \tilde{x}_1, \ldots, \tilde{x}_{\tilde{n}})$, pe baza căruia se trag concluzii (se fac *inferențe*) asupra parametrilor θ și θ_α și, eventual, asupra deciziilor care sînt influențate de acești parametrii, conform discuției din prima parte a lucrării.

Instrumentul principal pentru inferența statistică este furnizat de funcția $f(n, x_1, \ldots, x_n|\theta, \theta_\alpha)$, privită însă ca funcție de parametrii (θ, θ_α) în care $(n, x_1, \ldots, x_n)$ joacă rolul de parametrii. Funcția

$$L(\theta, \theta_\alpha|n, x_1, \ldots, x_n) = f(n, x_1, \ldots, x_n|\theta, \theta_\alpha) \tag{11.3.6}$$

este numită *funcția de verosimilitate* și este definită ori de cîte ori funcția din partea dreaptă a relației (11.3.6) este definită.

Exemplu. Pentru lămurirea noțiunilor introduse mai sus, să considerăm experimentul statistic următor.

Avem două monede. Probabilitatea de apariție a stemei într-o aruncare pentru prima monedă este $\theta \in (0, 1)$, iar probabilitatea de apariție a stemei într-o aruncare pentru a doua monedă este $\theta_\alpha \in (0, 1)$.

Scanned with OKEN Scanner

Experimentul se desfășoară în modul următor. Se aruncă mai întîi a doua monedă. Dacă apare stema ne oprim. Dacă apare valoarea, atunci aruncăm prima monedă și se notează rezultatul cu 1 dacă apare stema și cu 0 dacă apare valoarea. Se aruncă apoi a doua monedă: dacă apare stema ne oprim, iar dacă apare valoarea aruncăm din nou prima monedă ș.a.m.d.

Este clar că se pune în evidență un proces generator de date $\tilde{x}_1, \tilde{x}_2, \ldots$, în care variabilele aleatoare $\tilde{x}_i$ iau valoarea 1 sau 0 după cum la aruncarea de rang i a primei monede apare stema sau valoarea, șirul $(\tilde{x}_i)$ conținînd variabile aleatoare independente și identic repartizate cu funcția de masă

$$f(x|\theta) = \begin{cases} \theta & \text{dacă } x = 1, \\ 1-\theta & \text{dacă } x = 0, \\ 0 & \text{dacă } x \notin \{0, 1\}. \end{cases} \tag{11.3.7}$$

De asemenea, se pune în evidență un proces de oprire dat de probabilitățile de oprire

$$\pi(0|\theta, \theta_\alpha) = \pi(i|x_1, \ldots, x_i, \theta, \theta_\alpha) = \theta_\alpha \tag{11.3.8}$$

pentru orice i și orice $x_1, \ldots, x_i$, θ și θ_α.

Este clar că probabilitățile de a se face exact i observații sînt, conform formulelor (11.3.1)—(11.3.3), date de relațiile:

$$s(0|\theta, \theta_\alpha) = \theta_\alpha; \quad s(1|x_1, \theta, \theta_\alpha) = \theta_\alpha(1-\theta_\alpha)$$

și

$$s(i|x_1, \ldots, x_i, \theta, \theta_\alpha) = \theta_\alpha(1-\theta_\alpha)^i, \tag{11.3.9}$$

formulă care este valabilă, evident, pentru toți $i \geqslant 0$.

În sfîrșit, din formula (11.3.4), putem obține funcția de masă a vectorului aleator $(\tilde{n}, \tilde{x}_1, \ldots, \tilde{x}_{\tilde{n}})$:

$$f(n, x_1, \ldots, x_n|\theta, \theta_\alpha) = \theta_\alpha(1-\theta_\alpha)^n f(x_1, \ldots, x_n|\theta). \tag{11.3.10}$$

Ținînd însă seama de formula (11.3.5), rezultă că putem scrie

$$f(x_1, \ldots, x_n|\theta) = \begin{cases} \theta^r(1-\theta)^{n-r} & \text{dacă } r \in \{0, 1, \ldots, n\}, \\ 0 & \text{dacă } r \notin \{0, 1, \ldots, n\}, \end{cases}$$

Scanned with OKEN Scanner

unde $r = x_1 + \ldots + x_n$. Prin urmare putem scrie

$$f(n, x_1, \ldots, x_n | \theta, \theta_\alpha) =$$

$$= \begin{cases} \theta_\alpha(1-\theta_\alpha)^n \theta^r (1-\theta)^{n-r} & \text{dacă } r \in \{0, 1, \ldots, n\}, \\ 0 & \text{dacă } r \notin \{0, 1, \ldots, n\}, \end{cases} \quad (11.3.11)$$

unde $r = x_1 + \ldots + x_n$. Rezultă atunci că funcția de verosimilitate $L(\theta, \theta_\alpha | n, x_1, \ldots, x_n)$ este în acest caz funcție numai de parametrii n și $r = x_1 + \ldots + x_n$ și nu de toate rezultatele obținute.

11.4. INFERENȚĂ STATISTICĂ. STATISTICI SUFICIENTE

Inferența statistică se referă la concluziile care se pot trage asupra parametrilor $\boldsymbol{\theta}$ și $\boldsymbol{\theta}_\alpha$ ai unui experiment statistic pe baza rezultatului $(n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$ al acestui experiment. Cu alte cuvinte, după obținerea rezultatului $(n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$ al experimentului efectuat, trebuie luată o decizie în legătură cu (care este influențată de) parametrii $(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$. Din acest punct de vedere, inferența statistică poate fi privită ca un caz particular a problemei de decizie în condiții de incertitudine prezentată în prima parte a lucrării (vezi §1.3). Din motive ce vor deveni clare ceva mai tîrziu, este însă preferabil să tratăm această problemă separat.

Există două abordări principale ale problemelor de inferență statistică. Abordarea *clasică* sau *obiectivă* a problemei constă în utilizarea în scopul inferenței a funcției de verosimilitate $L(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha | n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$ dată de formula (11.3.6), considerată ca funcție de toate argumentele sale.

A doua abordare, pe care o vom numi *inferență bayesiană*, pleacă de la considerarea vectorului parametrilor ca un vector aleator $(\tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{\theta}}_\alpha)$, pentru care este cunoscută o *repartiție elementară inițială (apriori)* $f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$, care trebuie specificată. Concluziile sînt bazate în această abordare pe *repartiția elementară aposteriori* $f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha | n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$, condiționată de rezultatul $(n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$ al experimentului.

Scanned with OKEN Scanner

Față de abordarea bayesiană există unele rezerve și obiecții. Principala obiecție este următoarea: în multe probleme de statistică nu se poate specifica *în mod obiectiv* o repartiție elementară apriori $f(\theta, \theta_\alpha)$, chiar dacă procese de generare de date și procese de oprire pot fi definite în mod obiectiv pentru toate valorile posibile ale parametrilor (θ, θ_α).

Există preocuparea de a pune de acord cele două abordări, cel puțin în unele cazuri. Mai exact, se pot determina în unele cazuri repartiții elementare apriori $f(\theta, \theta_\alpha)$ cu proprietatea că, pentru fiecare rezultat posibil $(n, x_1, \ldots, x_n)$ al experimentului, repartiția elementară aposteriori $f(\theta, \theta_\alpha | n, x_1, \ldots, x_n)$ coincide cu funcția de verosimilitate $L(\theta, \theta_\alpha | n, x_1, \ldots, x_n)$. Pentru aceste repartiții elementare apriori, pe care le vom numi *proprii*, abordarea bayesiană și cea clasică sînt echivalente. Din acest punct de vedere, se poate spune că abordarea clasică este un caz particular al abordării bayesiene.

Există totuși cazuri în care nu există repartiții elementare apriori cu proprietatea specificată mai sus. În aceste cazuri, din egalarea funcției de verosimilitate cu repartiția elementară aposteriori rezultă funcții $f(\theta, \theta_\alpha)$ care nu satisfac condiția de normare

$$\int f(\theta, \theta_\alpha) = 1,$$

funcții care vor fi numite *repartiții elementare improprii*.

Indiferent de metoda de inferență pe care o utilizăm, concluziile pe care le tragem cu privire la parametrii (θ, θ_α) sînt bazate în mare parte, dacă nu în totalitate, pe rezultatul $(n, x_1, \ldots, x_n)$ al experimentului statistic efectuat. În general, datele obținute dintr-un experiment statistic sînt foarte numeroase, făcînd dificilă și, uneori, chiar imposibilă utilizarea lor ca atare în rezolvarea problemelor inferenței statistice. Este deci natural să încercăm o *condensare* a informației conținute în rezultatul experimentului $(n, x_1, \ldots, x_n)$. Această condensare a informației poate fi realizată cu ajutorul unor funcții măsurabile $T(., \ldots, .|n)$ definite pe $R^{k_1 + \cdots + k_n}$ și cu valori în R^{r_n}, unde k_i este numărul componentelor lui x_i, iar r_n este un număr întreg și pozitiv pentru fiecare n natural.

Definiția 11.4.1. *Numim statistică un șir de funcții măsurabile* $(T(., \ldots, .|n))_{n \geq 1}$, *cu proprietatea că pentru orice* $n = 1, 2, \ldots,$

funcția $T(., ..., .|n) : R^{k_1+...+k_n} \to R^{r_n}$ *este funcție numai de primele* n *valori observate* $x_1, ..., x_n$ *ale unui proces generator de date.*

Din această definiție rezultă că o statistică *nu depinde* de parametrii (θ, θ_α).

Remarcăm, de asemenea, că înainte de efectuarea unui experiment, n și $x_1, ..., x_n$ sînt variabilele aleatoare $\tilde{n}$ și $\tilde{x}_1, ..., \tilde{x}_n$, astfel încît $T(\tilde{x}_1, ..., \tilde{x}_n|\tilde{n})$ este de asemenea un vector aleator. Vom utiliza termenul de statistică referindu-ne fie la funcțiile $(T(., ..., .|n))_{n\geqslant 1}$, fie la vectorul aleator $\tilde{x}_T = T(\tilde{x}_1, ..., \tilde{x}_{\tilde{n}}|\tilde{n})$, fie la valoarea $x_T = T(x_1, ..., x_n|n)$.

Un exemplu imediat de statistică este furnizat de vectorul aleator $\tilde{x}_T = (\tilde{n}, \tilde{x}_1, ..., \tilde{x}_{\tilde{n}})$; este clar că această statistică nu micșorează volumul informației și deci nu realizează o condensare a informației conținută în rezultatul experimentului.

Un alt exemplu de statistică este $\tilde{x}_T = (\tilde{n}, \tilde{x}_1)$; această statistică micșorează în general volumul informației, dar nu realizează și o condensare a informației conținute în rezultatul experimentului.

Definiția 11.4.2. *O statistică* x_T *este suficientă pentru* (θ, θ_α) *dacă există funcțiile cu valori reale nenegative* $q(x_T|\theta, \theta_\alpha)$ *de* x_T *și având pe* θ *și* θ_α *drept parametri, și funcțiile cu valori reale nenegative* $q(n, x_1, ..., x_n)$, *astfel încît să avem*

$$f(n, x_1, ..., x_n|\theta, \theta_\alpha) = q(x_T|\theta, \theta_\alpha)\, q(n, x_1, ..., x_n) \quad (11.4.1)$$

pentru toate valorile posibile ale parametrilor (θ, θ_α) *și ale rezultatului experimentului* $(n, x_1, ..., x_n)$.

Din definiție rezultă că $q(n, x_1, ..., x_n)$ nu depinde de θ și θ_α, iar $q(x_T|\theta, \theta_\alpha)$ depinde de $(n, x_1, ..., x_n)$ numai prin intermediul statisticii x_T.

O definiție echivalentă este conținută în teorema următoare.

Teorema 11.4.1' *(Fisher-Neyman). Condiția necesară și suficientă ca* x_T *să fie o statistică suficientă pentru* (θ, θ_α) *este ca să existe o funcție* $q(x_T) > 0$ *astfel încît repartiția elementară a lui* $\tilde{x}_T$ *să fie dată de relația*

$$f(x_T|\theta, \theta_\alpha) = q(x_T|\theta, \theta_\alpha)\, q(x_T) \quad (11.4.2)$$

și repartiția elementară a lui $(\tilde{n}, \tilde{x}_1, \ldots, \tilde{x}_{\tilde{n}})$ *condiționată de* $\tilde{x}_T = x_T$ *să fie definită de*

$$f(n, x_1, \ldots, x_n | x_T, \theta, \theta_\alpha) = \frac{q(n, x_1, \ldots, x_n)}{q(x_T)}. \qquad (11.4.3)$$

Demonstrație. Necesitatea condițiilor (11.4.2) și (11.4.3) se deduce în modul următor. Deoarece x_T este o funcție numai de $n, x_1, \ldots, x_n$, rezulta că

$$f(n, x_1, \ldots, x_n | x_T, \theta, \theta_\alpha) = \frac{f(n, x_1, \ldots, x_n | \theta, \theta_\alpha)}{f(x_T | \theta, \theta_\alpha)} \qquad (11.4.4)$$

Funcția de repartiție a vectorului aleator $\tilde{x}_T$ poate fi scrisă sub forma

$$F(x_T | \theta, \theta_\alpha) = \int_{\Omega_n} f(n, x_1, \ldots, x_n | \theta, \theta_\alpha), \qquad (11.4.5)$$

unde

$$\Omega_n = \{(n, x_1, \ldots, x_n) : T(x_1, \ldots, x_n | n) < x_T\}. \qquad (11.4.6)$$

Ținînd seama de (11.4.1) rezultă că (11.4.5) poate fi scris sub forma

$$F(x_T | \theta, \theta_\alpha) = q(x_T | \theta, \theta_\alpha) \int_{\Omega_n} q(n, x_1, \ldots, x_n),$$

ceea ce arată că repartiția elementară a lui $\tilde{x}_T$ este de forma

$$f(x_T | \theta, \theta_\alpha) = q(x_T | \theta, \theta_\alpha)\, q(x_T) \qquad (11.4.7)$$

cu $q(x_T) > 0$. Prin urmare (11.4.2) este obținută.

Din (11.4.1), (11.4.4) și (11.4.7) rezultă acum

$$f(n, x_1, \ldots, x_n | x_T, \theta, \theta_\alpha) = \frac{q(x_T | \theta, \theta_\alpha)\, q(n, x_1, \ldots, x_n)}{q(x_T | \theta, \theta_\alpha)\, q(x_T)},$$

de unde se obține (11.4.3).

Scanned with OKEN Scanner

Suficiența condițiilor (11.4.2) și (11.4.3) se deduce de asemenea ușor. Într-adevăr, utilizînd din nou relația (11.4.4), obținem

$$f(n, \mathbf{x}_1, \ldots, \mathbf{x}_n|\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha) = f(\mathbf{x}_T|\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)\, f(n, \mathbf{x}_1, \ldots, \mathbf{x}_n|\mathbf{x}_T, \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$$

și, ținînd seama de (11.4.2) și (11.4.3), putem scrie

$$f(n, \mathbf{x}_1, \ldots, \mathbf{x}_n|\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha) = q(\mathbf{x}_T|\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)\, q(\mathbf{x}_T)\, \frac{q(n, \mathbf{x}_1, \ldots, \mathbf{x}_n)}{q(\mathbf{x}_T)},$$

de unde rezultă imediat (11.4.1).

Observație. Condiția (11.4.3) ne arată că repartiția elementară a vectorului aleator $(\tilde{n}, \tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_{\tilde{n}})$ condiționată de $\tilde{\mathbf{x}}_T$ este independentă de $(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$ și deci putem scrie $f(n, \mathbf{x}_1, \ldots, \mathbf{x}_n|\mathbf{x}_T)$ în loc de $f(n, \mathbf{x}_1, \ldots, \mathbf{x}_n|\mathbf{x}_T, \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$. Cu această observație și utilizînd (11.4.2) și (11.4.3), rezultă că (11.4.1) poate fi scrisă sub forma

$$f(n, \mathbf{x}_1, \ldots, \mathbf{x}_n|\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha) = f(\mathbf{x}_T|\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)\, f(n, \mathbf{x}_1, \ldots, \mathbf{x}_n|\mathbf{x}_T). \quad (11.4.8)$$

Rezultă deci că putem defini o statistică suficientă $\mathbf{x}_T$ pentru $(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$ ca fiind statistica cu proprietatea (11.4.8). Definiția a fost dată în forma (11.4.1) deoarece, în general, este mai ușor de verificat (11.4.1) decît (11.4.8).

Un rezultat fundamental privind statisticile suficiente este următorul.

Teorema 11.4.2. *Fie* $\mathrm{f}(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$ *o repartiție elementară apriori pentru vectorul aleator* $(\tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{\theta}}_\alpha)$ *și fie* $\mathbf{x}_\mathrm{T}$ *o statistică suficientă pentru* $(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$.

Dacă $\mathrm{f}(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha|\mathbf{x}_\mathrm{T})$ *și* $\mathrm{f}(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha|\mathrm{n}, \mathbf{x}_1, \ldots, \mathbf{x}_\mathrm{n})$ *sînt repartițiile elementare aposteriori corespunzătoare lui* $\mathrm{f}(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$, *atunci*

$$\mathrm{f}(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha|\mathbf{x}_\mathrm{T}) = \mathrm{f}(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha|\mathrm{n}, \mathbf{x}_1, \ldots, \mathbf{x}_\mathrm{n}). \quad (11.4.9)$$

Demonstrație. Aplicînd teorema lui Bayes, putem scrie

$$f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha|n, \mathbf{x}_1, \ldots, \mathbf{x}_n) = \frac{f(n, \mathbf{x}_1, \ldots, \mathbf{x}_n|\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)\, f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)}{\iint f(n, \mathbf{x}_1, \ldots, \mathbf{x}_n|\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)\, f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)\, \mathrm{d}\boldsymbol{\theta}\, \mathrm{d}\boldsymbol{\theta}_\alpha}$$

unde integrala se face în raport cu toate valorile posibile ale parametrilor $(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$.

Scanned with OKEN Scanner

Utilizînd (11.4.8) putem scrie

$$f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha | n, \mathbf{x}_1, \ldots, \mathbf{x}_n) =$$

$$= \frac{f(\mathbf{x}_T | \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)\, f(n, \mathbf{x}_1, \ldots, \mathbf{x}_n | \mathbf{x}_T)\, f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)}{f(n, \mathbf{x}_1, \ldots, \mathbf{x}_n | \mathbf{x}_T) \iint f(\mathbf{x}_T | \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)\, f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)\, d\boldsymbol{\theta}\, d\boldsymbol{\theta}_\alpha}$$

dacă ținem seama că $f(n, \mathbf{x}_1, \ldots, \mathbf{x}_n | \mathbf{x}_T)$ este o funcție numai de $(n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$. Rezultă atunci că

$$f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha | n, \mathbf{x}_1, \ldots, \mathbf{x}_n) = \frac{f(\mathbf{x}_T | \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)\, f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)}{\iint f(\mathbf{x}_T | \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)\, f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)\, d\boldsymbol{\theta}\, d\boldsymbol{\theta}_\alpha} = f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha | \mathbf{x}_T),$$

ceea ce demonstrează (11.4.9).

Menționăm, fără demonstrație, faptul că reciproca teoremei 11.4.2 este de asemenea adevărată.

Teorema 11.4.3. *Fie* $\mathbf{x}_T$ *o statistică. Dacă pentru orice repartiție elementară apriori* $f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$ *avem*

$$f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha | \mathbf{x}_T) = f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha | n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$$

pentru toate valorile posibile ale parametrilor $(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$ *și ale valorilor observate* $(n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$, *atunci* $\mathbf{x}_T$ *este o statistică suficientă pentru* $(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$.

Cititorul interesat poate găsi o demonstrație a acestei teoreme în Raiffa, H, Schlaifer, R. (1961).

De remarcat că pentru luarea deciziei este posibil să fie relevant numai un anumit subvector $\boldsymbol{\theta}_1$ al vectorului parametrilor $(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$. De fapt, în mod obișnuit, $\boldsymbol{\theta}_1$ este un subvector al lui $\boldsymbol{\theta}$, vector care conține parametrii procesului de generare de date; în unele cazuri, însă, este posibil ca $\boldsymbol{\theta}_1$ să conțină și componente ale lui $\boldsymbol{\theta}_\alpha$, din cauza influenței pe care poate s-o aibă procesul de oprire al experimentului pentru inferența asupra lui $\boldsymbol{\theta}$. Dacă notăm prin $\boldsymbol{\theta}_2$ vectorul format cu componentele lui $(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$ care nu se găsesc în $\boldsymbol{\theta}_1$, rezultă că $(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$ poate fi partiționat în $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$.

Scanned with OKEN Scanner

Din demonstrația teoremei 11.4.2 rezultă ușor că dacă $\mathbf{x}_T$ este o statistică suficientă pentru $(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$, atunci

$$f(\boldsymbol{\theta}_1|\mathbf{x}_T) = f(\boldsymbol{\theta}_1|n, \mathbf{x}_1, \ldots, \mathbf{x}_n) \tag{11.4.10}$$

pentru toate valorile lui $\boldsymbol{\theta}_1$ și pentru toate valorile posibile $(n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$. Relația (11.4.10) arată că repartițiile elementare aposteriori pentru $\boldsymbol{\theta}_1$ condiționate de $\mathbf{x}_T$ și respectiv de $(n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$ sînt egale. Cu alte cuvinte, dacă $\mathbf{x}_T$ este suficientă pentru $(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$, atunci $\mathbf{x}_T$ este de asemenea suficientă pentru $\boldsymbol{\theta}_1$. Așadar, o statistică suficientă este *marginal suficientă.*

De obicei, repartițiile apriori sînt restrînse la o anumită clasă precizată. Fie $(f_i(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha))_{i \in I}$ o familie fixată de repartiții elementare apriori pentru $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$, unde I este o mulțime arbitrară. Vom nota prin $f_i(\boldsymbol{\theta}_1|n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$ și $f_i(\boldsymbol{\theta}_1|\mathbf{x}_T)$ repartițiile elementare marginale aposteriori condiționate respectiv de $(n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$ și $\mathbf{x}_T$.

Definiția 11.4.3. *Spunem că statistica* $\mathbf{x}_T$ *este marginal suficientă pentru* $\boldsymbol{\theta}_1$, *relativ la familia*

$$\mathfrak{L} = (\mathrm{f}_i(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2))_{i \in I}$$

dacă

$$\mathrm{f}_i(\boldsymbol{\theta}_1|\mathbf{x}_T) = \mathrm{f}_i(\boldsymbol{\theta}_1|\mathrm{n}, \mathbf{x}_1, \ldots, \mathbf{x}_n) \tag{11.4.11}$$

pentru orice $i \in I$, *pentru orice* $\boldsymbol{\theta}_1$ *și pentru orice valori* $(n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$ *pentru care* $\mathrm{f}_i(n, \mathbf{x}_1, \ldots, \mathbf{x}_n) > 0$ *pentru toți* $i \in I$.

Este clar că dacă $\mathbf{x}_T$ este o statistică marginal suficientă pentru $\boldsymbol{\theta}_1$ relativ la o familie de repartiții elementare apriori care conține repartiția elementară inițială care interesează, atunci, chiar dacă $\mathbf{x}_T$ nu este suficientă pentru $(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$, singurul lucru pe care trebuie să-l reținem este această statistică.

Componentele vectorului $\boldsymbol{\theta}_2$, care nu au legătură directă cu problema de decizie, dar care trebuie luate în considerare în analiză, sînt numite uneori *parametrii inconvenienți.*

Dacă familia de repartiții elementare apriori este de forma*

$$\mathfrak{L} = \{f(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) : f(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = f(\boldsymbol{\theta}_1) f(\boldsymbol{\theta}_2) \text{ pentru orice } (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)\}, \tag{11.4.12}$$

* Rezultă că vectorii $\boldsymbol{\theta}_1$ și $\boldsymbol{\theta}_2$ sînt independenți.

Scanned with OKEN Scanner

atunci se poate obține o condiție echivalentă de marginal suficiență relativă la $\mathfrak{L}$ a unei statistici. Mai precis, avem rezultatul următor.

Teorema 11.4.4. *Statistica* $\mathbf{x}_T$ *este marginal suficientă pentru* θ_1 *relativ la familia* $\mathfrak{L}$ *dată de* (11.4.12) *dacă și numai dacă există funcțiile* $q(\mathbf{x}_T|\theta_1) \geqslant 0$ *și* $q(n, \mathbf{x}_1, \ldots, \mathbf{x}_n|\theta_2) \geqslant 0$ *astfel încît*

$$f(n, \mathbf{x}_1, \ldots, \mathbf{x}_n|\theta_1, \theta_2) = \\ = q(\mathbf{x}_T|\theta_1)\, q(n, \mathbf{x}_1, \ldots, \mathbf{x}_n|\theta_2), \tag{11.4.13}$$

ori de cîte ori membrul stîng este definit.

Demonstrația acestei teoreme nu comportă dificultăți, raționamentul fiind de aceeași natură cu cel utilizat în demonstrația teoremei 11.4.1.

În sfîrșit, după cum am avut ocazia să menționăm (vezi formula (11.3.4)), pentru a calcula repartiția elementară marginală aposteriori $f(\theta|n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$, este necesar să utilizăm atît repartiția elementară $f(n, \mathbf{x}_1, \ldots, \mathbf{x}_n|\theta)$, cît și probabilitățile $s(n|\mathbf{x}_1, \ldots, \mathbf{x}_n, \theta, \theta_\alpha)$ ca experimentul să se oprească la etapa n (sau, echivalent, probabilitatea ca observația $n+1$ să nu se efectueze). Teorema următoare dă condiții suficiente ca în calculul lui $f(\theta|n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$ să utilizăm numai $f(\mathbf{x}_1, \ldots, \mathbf{x}_n|\theta)$ și repartiția elementară apriori $f(\theta)$, fără utilizarea probabilităților de oprire $s(n|\mathbf{x}_1, \ldots, \mathbf{x}_n, \theta, \theta_\alpha)$, al căror calcul poate fi dificil în unele cazuri.

Teorema 11.4.5. *Să presupunem că sînt îndeplinite condițiile următoare:*

(a) Pentru orice n, *probabilitățile* $\pi(n|\mathbf{x}_1, \ldots, \mathbf{x}_n, \theta, \theta_\alpha)$ *nu depind de* θ;

(b) Repartiția elementară apriori aparține clasei

$$\mathfrak{L}_1 = \{f(\theta, \theta_\alpha) : f(\theta, \theta_\alpha) = f(\theta)\, f(\theta_\alpha) \text{ pentru orice } (\theta, \theta_\alpha)\}. \tag{11.4.14}$$

În aceste condiții avem

$$f(\theta|n, \mathbf{x}_1, \ldots \mathbf{x}_n) = \frac{f(\mathbf{x}_1, \ldots, \mathbf{x}_n|\theta)\, f(\theta)}{\int f(\mathbf{x}_1, \ldots, \mathbf{x}_n|\theta)\, f(\theta)\, d\theta} \tag{11.4.15}$$

ori de cîte ori termenii egalității sînt definiți.

Scanned with OKEN Scanner

Demonstrație. Din condiția (*a*) rezultă că $s(n|\mathbf{x}_1, \ldots, \mathbf{x}_n, \boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)$ sînt de asemenea independente de $\boldsymbol{\theta}$ pentru orice n.

Este clar că

$$f(\boldsymbol{\theta}|n, \mathbf{x}_1, \ldots, \mathbf{x}_n) = \int f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha|n, \mathbf{x}_1, \ldots, \mathbf{x}_n)\, \mathrm{d}\boldsymbol{\theta}_\alpha, \qquad (11.4.16)$$

unde integrala se face în raport cu $\boldsymbol{\theta}_\alpha$, și

$$f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha|n, \mathbf{x}_1, \ldots, \mathbf{x}_n) = \frac{f(n, \mathbf{x}_1, \ldots, \mathbf{x}_n|\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)\, f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)}{\iint f(n, \mathbf{x}_1, \ldots, \mathbf{x}_n|\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)\, f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha)\, \mathrm{d}\boldsymbol{\theta}\, \mathrm{d}\boldsymbol{\theta}_\alpha}.$$

Utilizînd acum (11.3.4) rezultă că

$$f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha|\, n, \mathbf{x}_1, \ldots, \mathbf{x}_n) = \frac{s(n|\mathbf{x}_1, \ldots, \mathbf{x}_n, \boldsymbol{\theta}_\alpha)\, f(\mathbf{x}_1, \ldots, \mathbf{x}_n|\boldsymbol{\theta})\, f(\boldsymbol{\theta})\, f(\boldsymbol{\theta}_\alpha)}{\iint s(n|\mathbf{x}_1, \ldots, \mathbf{x}_n, \boldsymbol{\theta}_\alpha)\, f(\mathbf{x}_1, \ldots, \mathbf{x}_n|\boldsymbol{\theta})\, f(\boldsymbol{\theta})\, f(\boldsymbol{\theta}_\alpha)\, \mathrm{d}\boldsymbol{\theta}\, \mathrm{d}\boldsymbol{\theta}_\alpha}.$$

sau

$$f(\boldsymbol{\theta}, \boldsymbol{\theta}_\alpha|n, \mathbf{x}_1, \ldots, \mathbf{x}_n) = \frac{s(n|\mathbf{x}_1, \ldots, \mathbf{x}_n, \boldsymbol{\theta}_\alpha)\, f(\mathbf{x}_1, \ldots, \mathbf{x}_n|\boldsymbol{\theta})\, f(\boldsymbol{\theta})\, f(\boldsymbol{\theta}_\alpha)}{\left[\int s(n|\mathbf{x}_1, \ldots, \mathbf{x}_n, \boldsymbol{\theta}_\alpha)\, f(\boldsymbol{\theta}_\alpha)\, \mathrm{d}\boldsymbol{\theta}_\alpha\right]\left[\int f(\mathbf{x}_1, \ldots, \mathbf{x}_n|\boldsymbol{\theta})\, f(\boldsymbol{\theta})\, \mathrm{d}\boldsymbol{\theta}\right]} \qquad (11.4.17)$$

Utilizînd acum (11.4.16) și (11.4.17) obținem ușor (11.4.15) și teorema este astfel demonstrată.

Observație. Teorema arată că, în condițiile (*a*), (*b*), pentru inferența bayesiană putem considera experimentul ca și cînd numărul n al observațiilor ar fi fost predeterminat. Vom spune în acest caz că procesul de oprire este *neinformativ* asupra lui $\boldsymbol{\theta}$, cazul contrar fiind numit proces de oprire *informativ* asupra lui $\boldsymbol{\theta}$. La procesul de oprire neinformativ ne vom referi de asemenea prin expresia „obținerea a n observații $\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_n$

Scanned with OKEN Scanner

dintr-un proces generator de date cu oprire neinformativă" sau, încă, prin expresia „$\tilde{x}_1, \ldots, \tilde{x}_n$ este o selecție (aleatoare) de volum n din $f(\mathbf{x}|\boldsymbol{\theta})$".

Vom încheia aceste considerații succinte privind inferența statistică prin prezentarea în linii generale a trei probleme fundamentale ale inferenței statistice: estimarea punctuală, estimarea prin regiuni de încredere și testarea ipotezelor statistice.

Prin estimare punctuală vom înțelege problema care constă în determinarea unei estimații „bune" $\hat{\boldsymbol{\theta}}$ pentru $\boldsymbol{\theta}$, estimație care să fie bazată pe rezultatul experimentului $(n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$. Cuvîntul „bun" trebuie evident precizat, atît în inferența clasică cît și în cea bayesiană existînd mai multe moduri de a interpreta acest cuvînt.

Prin estimare prin regiuni de încredere înțelegem problema furnizării unei regiuni „destul de mici" $D(n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$ care să aibă proprietatea că „este rezonabil să se presupună" că $\boldsymbol{\theta}$ aparține lui $D(n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$. Evident, cuvintele puse între ghilimele trebuie precizate, atît în inferența clasică cît și în cea bayesiană. Dacă $\boldsymbol{\theta}$ este un scalar, $D(n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$ este numit *interval de încredere* pentru $\boldsymbol{\theta}$. Este clar că estimarea prin regiuni de încredere este o problemă mai puțin ambițioasă decît problema estimării punctuale.

În sfîrșit, a treia problemă menționată constă în acceptarea sau respingerea ipotezei (afirmației) că $\boldsymbol{\theta}$ aparține unei anumite submulțimi Θ_0 a mulțimii Θ a tuturor valorilor lui $\boldsymbol{\theta}$, acceptare sau respingere care se bazează pe rezultatul $(n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$ al unui experiment statistic. Între testarea ipotezelor statistice și estimarea prin regiuni de încredere există o legătură strînsă, care poate fi constatată chiar din prezentarea generală făcută mai sus celor două probleme.

11.5. STATISTICI UZUALE

În rezolvarea problemelor de inferență statistică prezentate mai sus se utilizează pe scară largă o gamă variată de statistici, cele mai utilizate fiind construite pe baza așa-numitelor *momente de selecție*.

Scanned with OKEN Scanner

Dacă $\tilde{x}_1, \tilde{x}_2, \ldots$ este un proces generator de date unidimensional de parametru $\mathbf{0}$ și se utilizează un proces de oprire neinformativ asupra lui $\mathbf{0}$, atunci vom introduce *momentele de selecție* ca fiind variabilele aleatoare

$$\tilde{m}_r = \frac{1}{n} \sum_{i=1}^{n} \tilde{x}_i^r. \tag{11.5.1}$$

Momentul de selecție $\tilde{m}_1$ este numit și *medie de selecție* și este notat prin $\tilde{\bar{x}}$, adică

$$\tilde{\bar{x}} = \frac{1}{n} \sum_{i=1}^{n} \tilde{x}_i. \tag{11.5.2}$$

Analog, numim *momente centrate de selecție*, variabilele aleatoare

$$\tilde{v}_r = \frac{1}{n} \sum_{i=1}^{n} (\tilde{x}_i - \tilde{\bar{x}})^r. \tag{11.5.3}$$

Momentul centrat de selecție $\tilde{v}_2$ este numit și *dispersie de selecție*.

Este clar că valoarea medie a unui moment de selecție este egală cu momentul de același ordin al variabilelor aleatoare $(\tilde{x}_i)$, adică

$$M(\tilde{m}_r) = M(\tilde{x}_i^r), \quad 1 \leqslant i \leqslant n, \tag{11.5.4}$$

iar dispersia momentelor de ordinul r este de forma

$$V(\tilde{m}_r) = \frac{1}{n} [M(\tilde{x}_i^{2r}) - M^2(\tilde{x}_i^r)], \quad 1 \leqslant i \leqslant n, \tag{11.5.5}$$

cum se vede imediat. În particular, pentru $r = 1$, obținem

$$M(\tilde{\bar{x}}) = M(\tilde{x}_i), \quad 1 \leqslant i \leqslant n, \tag{11.5.6}$$

și

$$V(\tilde{\bar{x}}) = \frac{1}{n} V(\tilde{x}_i), \quad 1 \leqslant i \leqslant n. \tag{11.5.7}$$

Exprimarea valorii medii și a dispersiei pentru momentele centrate de selecție este ceva mai complicată. Din acest motiv, ne vom limita la cazul particular $r = 2$, adică la dispersia de selecție $\tilde{v}_2$, care este utilizată în mod frecvent.

Scanned with OKEN Scanner

Pentru calculul valorii medii a lui $\tilde{v}_2$ este util să observăm că putem scrie

$$\tilde{v}_2 = \frac{1}{n}\sum_{i=1}^{n}(\tilde{x}_i^2 - 2\tilde{x}_i\tilde{\bar{x}} + \tilde{\bar{x}}^2) = \frac{1}{n}\sum_{i=1}^{n}\tilde{x}_i^2 - (\tilde{\bar{x}})^2$$

sau

$$\tilde{v}_2 = \frac{1}{n}\sum_{i=1}^{n}\tilde{x}_i^2 - \frac{1}{n^2}\sum_{i=1}^{n}\tilde{x}_i^2 - \frac{1}{n^2}\sum_{i\neq j}\tilde{x}_i\tilde{x}_j.$$

Trecînd la valoarea medie și ținînd seama că $\tilde{x}_1, \ldots, \tilde{x}_n$ sînt independente și identic repartizate, obținem

$$M(\tilde{v}_2) = \left(\frac{1}{n} - \frac{1}{n^2}\right) nM(\tilde{x}_i^2) - \frac{n(n-1)}{n^2}M^2(\tilde{x}_i),$$

adică

$$M(\tilde{v}_2) = \frac{n-1}{n}V(\tilde{x}_i), \quad 1 \leqslant i \leqslant n. \tag{11.5.8}$$

Pentru calculul dispersiei lui $\tilde{v}_2$ vom observa mai întîi că punînd

$$\tilde{y}_i = \tilde{x}_i - M(\tilde{x}_i), \quad 1 \leqslant i \leqslant n,$$

obținem

$$\tilde{v}_2 = \frac{1}{n}\sum_{i=1}^{n}(\tilde{x}_i - \tilde{\bar{x}})^2 = \frac{1}{n}\sum_{i=1}^{n}(\tilde{y}_i - \tilde{\bar{y}})^2,$$

unde $M(\tilde{y}_i) = 0$, $1 \leqslant i \leqslant n$. Cu alte cuvinte, putem presupune, fără a restrînge generalitatea, că $\tilde{x}_1, \ldots, \tilde{x}_n$ au media zero și deci momentele lor coincid cu momentele centrate.

Rezultă prin urmare că putem scrie

$$\tilde{v}_2^2 = \frac{1}{n^2}\left(\sum_{i=1}^{n}\tilde{x}_i^2 - n(\tilde{\bar{x}})^2\right)^2$$

sau

$$\tilde{v}_2^2 = \frac{1}{n^2}\left[\left(\sum_{i=1}^{n}\tilde{x}_i^2\right)^2 - 2n(\tilde{\bar{x}})^2\sum_{i=1}^{n}\tilde{x}_i^2 + n^2(\tilde{\bar{x}})^4\right]. \tag{11.5.9}$$

Scanned with OKEN Scanner

Trecînd la valoarea medie obținem

$$M(\tilde{v}_2^2) = \frac{1}{n^2} M\left[\left(\sum_{i=1}^{n} \tilde{x}_i^2\right)^2\right] - \frac{2}{n} M\left[(\tilde{\bar{x}})^2 \sum_{i=1}^{n} \tilde{x}_i^2\right] + M((\tilde{\bar{x}})^4). \tag{11.5.10}$$

În continuare vom evalua succesiv cei trei termeni care apar în membrul drept al relației (11.5.10). Avem mai întîi:

$$\frac{1}{n^2} M\left(\left(\sum_{i=1}^{n} \tilde{x}_i^2\right)^2\right) = \frac{1}{n^2} M\left(\sum_{i=1}^{n} \tilde{x}_i^4 + \sum_{i \neq j} \tilde{x}_i^2 \tilde{x}_j^2\right) =$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} M(\tilde{x}_i^4) + \frac{1}{n^2} \sum_{i \neq j} M(\tilde{x}_i^2)\, M(\tilde{x}_j^2) =$$

$$= \frac{1}{n} M(\tilde{x}_i^4) + \frac{n-1}{n} (V(\tilde{x}_i))^2. \tag{11.5.11}$$

Pentru al doilea termen obținem evaluarea

$$\frac{2}{n} M\left[(\tilde{\bar{x}})^2 \sum_{i=1}^{n} \tilde{x}_i^2\right] = \frac{2}{n^3} M\left(\sum_{i,j=1}^{n} \tilde{x}_i \tilde{x}_j \sum_{k=1}^{n} \tilde{x}_k^2\right) =$$

$$= \frac{2}{n^3} M\left(\sum_{i=1}^{n} \tilde{x}_i^4 + \sum_{i \neq j} \tilde{x}_i^2 \tilde{x}_j^2 + \sum_{i \neq j} \tilde{x}_i \tilde{x}_j \sum_{\substack{k \neq i \\ k \neq j}} \tilde{x}_k^2\right) =$$

$$= \frac{2n}{n^3} M(\tilde{x}_i^4) + \frac{2n(n-1)}{n^3} M(\tilde{x}_i^2)\, M(\tilde{x}_j^2) + 0 =$$

$$= \frac{2}{n^2} [M(\tilde{x}_i^4) + (n-1)\,(V(\tilde{x}_i))^2]. \tag{11.5.12}$$

În sfîrșit, pentru al treilea termen din (11.5.10) avem

$$M(\tilde{\bar{x}}^4) = \frac{1}{n^4} M\left(\sum_{i=1}^{n} x_i^4 + 3 \sum_{i \neq j} \tilde{x}_i^2 \tilde{x}_j^2 + \ldots\right) =$$

$$= \frac{1}{n^4} [nM(\tilde{x}_i^4) + 3n(n-1)\,(V(\tilde{x}_i))^2] =$$

Scanned with OKEN Scanner

$$= \frac{1}{n^3}[M(\tilde{x}_i^4) + 3(n-1)\,(V(\tilde{x}_i))^2]; \qquad (11.5.13)$$

termenii nescriși din prima paranteză conțin pe $\tilde{x}_i$ și deci au media nulă.

Din relațiile (11.5.10)—(11.5.13) se obține ușor

$$M(\tilde{v}_2^2) = \frac{n-1}{n^3}[(n-1)\,M(\tilde{x}_i^4) + (n^2 - 2n + 3)\,(V(\tilde{x}_i))^2] \qquad (11.5.14)$$

pentru $1 \leqslant i \leqslant n$.

Deoarece dispersia lui $\tilde{v}_2$ se scrie sub forma

$$V(\tilde{v}_2) = M(\tilde{v}_2^2) - M^2(\tilde{v}_2),$$

din (11.5.8) și (11.5.14) rezultă imediat

$$V(\tilde{v}_2) = \frac{n-1}{n^3}[(n-1)\,M(\tilde{x}_i^4) - (n-3)\,(V(\tilde{x}_i))^2]. \qquad (11.5.15)$$

Să observăm de asemenea că statistica

$$\tilde{s}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\tilde{x}_i - \tilde{\bar{x}})^2 \qquad (11.5.16)$$

sau

$$\tilde{s}^2 = \frac{n}{n-1}\,\tilde{v}_2 \qquad (11.5.17)$$

are media și dispersia date de relațiile:

$$M(\tilde{s}^2) = V(\tilde{x}_i), \quad 1 \leqslant i \leqslant n; \qquad (11.5.18)$$

$$V(\tilde{s}^2) = \frac{1}{n}\left[M(\tilde{x}_i^4) - \frac{n-3}{n-1}\,(V(\tilde{x}_i))^2\right], \quad 1 \leqslant i \leqslant n. \qquad (11.5.19)$$

Scanned with OKEN Scanner

11.6. PROCESUL NORMAL UNIDIMENSIONAL

11.6.1. INTRODUCERE

În § 10.2 am introdus repartiția normală, pe care am studiat-o amănunțit, iar în § 10.5 am arătat rolul central pe care această repartiție îl joacă în studiul multor fenomene aleatoare.

În cele ce urmează vom considera experimente statistice în care apar *procese normale* (sau *gaussiene*) de generare de date. Mai precis, vom spune că un șir finit sau infinit de variabile aleatoare $(\tilde{x}_i)$ este un *proces normal* de generare de date de parametrii m și $\sigma^2 > 0$ dacă variabilele aleatoare $(\tilde{x}_i)$ sînt independente și identic repartizate și aparțin clasei $N(m, \sigma^2)$. Vom spune, de asemenea, că $\tilde{x}_1, \ldots, \tilde{x}_n$ reprezintă *o selecție de volum n* din populația asociată unei variabile aleatoare $\tilde{x} \in N(m, \sigma^2)$.

Este clar că, relativ la cei doi parametri m și σ^2, putem considera următoarele cazuri posibile*:

Cazul A: m este necunoscut, iar σ^2 este cunoscut;
Cazul B: m este cunoscut, iar σ^2 este necunoscut;
Cazul C: m și σ^2 sînt necunoscuți.

În cele ce urmează vom considera un proces de oprire neinformativ, în mod obișnuit cel caracterizat prin fixarea în prealabil a numărului n de observații. Pentru un proces de oprire neinformativ arbitrar, vom considera statisticile

$$\tilde{\bar{x}} = \frac{1}{n} \sum_{i=1}^{n} \tilde{x}_i \tag{11.6.1}$$

și

$$\tilde{s}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\tilde{x}_i - \tilde{\bar{x}})^2, \tag{11.6.2}$$

introduse în § 11.5.

În cazul (B) vom considera și statistica

$$\tilde{s}^2(m) = \frac{1}{n} \sum_{i=1}^{n} (\tilde{x}_i - m)^2. \tag{11.6.3}$$

---

* Cazul cînd m și σ^2 sînt cunoscuți este banal din punctul de vedere al inferenței statistice.

Din rezultatele generale stabilite în § 11.5 rezultă imediat că dacă variabilele aleatoare $\tilde{x}_1, \ldots, \tilde{x}_n$ constituie o selecție de volum n din populația asociată variabilei aleatoare $\tilde{x} \in N(m, \sigma^2)$, atunci avem:

$$M(\tilde{\bar{x}}) = m; \quad V(\tilde{\bar{x}}) = \frac{1}{n}\sigma^2 \tag{11.6.4}$$

și

$$M(\tilde{s}^2) = \sigma^2; \quad V(\tilde{s}^2) = \frac{1}{n}\left(M(\tilde{x}^4) - \frac{n-3}{n-1}\sigma^4\right). \tag{11.6.5}$$

Dacă ținem seama de faptul că momentul centrat de ordinul patru este $3\sigma^4$ (vezi formula (10.3.12)), atunci se obține imediat

$$V(\tilde{s}^2) = \frac{2\sigma^4}{n-1} + \frac{m^2(6\sigma^2 + m^2)}{n}. \tag{11.6.6}$$

Dacă m este cunoscut se poate observa că

$$n\tilde{s}^2(m) = n(\tilde{\bar{x}} - m)^2 + (n-1)\tilde{s}^2. \tag{11.6.7}$$

Într-adevăr, se poate scrie

$$n\tilde{s}^2(m) = \sum_{i=1}^{n}(\tilde{x}_i - m)^2 = \sum_{i=1}^{n}(\tilde{x}_i - \tilde{\bar{x}} + \tilde{\bar{x}} - m)^2,$$

adică

$$n\tilde{s}^2(m) = n(\tilde{\bar{x}} - m)^2 + 2(\tilde{\bar{x}} - m)\sum_{i=1}^{n}(\tilde{x}_i - \tilde{\bar{x}}) + \sum_{i=1}^{n}(\tilde{x}_i - \tilde{\bar{x}})^2.$$

Deoarece

$$\sum_{i=1}^{n}(\tilde{x}_i - \tilde{\bar{x}}) = 0, \tag{11.6.8}$$

formula (11.6.7) se deduce imediat din ultima relație obținută mai sus.

Scanned with OKEN Scanner

Luînd valoarea medie în (11.6.7) și avînd în vedere rezultatele anterioare, obținem

$$M(\tilde{s}^2(m)) = V(\tilde{\bar{x}}) + \frac{n-1}{n} M(\tilde{s}^2), \tag{11.6.9}$$

de unde

$$M(\tilde{s}^2(m)) = \sigma^2. \tag{11.6.10}$$

Pentru a determina dispersia lui $\tilde{s}^2(m)$, să observăm că

$$\tilde{s}^4(m) = \frac{1}{n^2}\left[\sum_{i=1}^{n} (\tilde{x}_i - m)^2\right]^2 =$$

$$= \frac{1}{n^2}\left[\sum_{i=1}^{n} (\tilde{x}_i - m)^4 + \sum_{i \neq j} (\tilde{x}_i - m)^2 (\tilde{x}_j - m)^2\right]$$

și prin urmare

$$M(\tilde{s}^4(m)) = \frac{1}{n^2}\left[n \cdot 3\sigma^4 + n(n-1)\sigma^4\right] = \frac{n+2}{n}\sigma^4.$$

Rezultă atunci că

$$V(\tilde{s}^2(m)) = M(\tilde{s}^4(m)) - M^2(\tilde{s}^2(m)) = \frac{n+2}{n}\sigma^4 - \sigma^4,$$

adică

$$V(\tilde{s}^2(m)) = \frac{2\sigma^4}{n}. \tag{11.6.11}$$

În paragraful următor vom studia procesul normal de generare de date în fiecare dintre cele trei cazuri menționate mai sus, prezentînd de fiecare dată funcția de verosimilitate asociată.

11.6.2. FUNCȚII DE VEROSIMILITATE

Vom prezenta forma funcțiilor de verosimilitate pentru un proces $(\tilde{x}_i)$ de generare de date normal cu parametrii m și σ^2 în cazul cînd procesul de oprire este neinformativ asupra parametrului în discuție.

Scanned with OKEN Scanner

Teorema 11.6.1. *În cazul cînd* m *este necunoscut, iar* $\sigma^2 > 0$ *este cunoscut, funcția de verosimilitate pentru* m *este*

$$L(m \mid n, x_1, \ldots, x_n, \sigma^2) = k_1 \exp\left[-\frac{1}{2} n(\bar{x} - m)^2/\sigma^2 \right], \tag{11.6.12}$$

unde

$$k_1 = (2\pi\sigma^2)^{-n/2} \exp\left[-\frac{1}{2}(n-1)s^2/\sigma^2 \right]. \tag{11.6.13}$$

În plus, statistica $(n, \tilde{\bar{x}})$ *este suficientă pentru* m.

Demonstrație. Funcția de verosimilitate pentru m este prin definiție

$$L(m \mid n, x_1, \ldots, x_n, \sigma^2) = f(x_1, \ldots, x_n \mid m, \sigma^2) =$$

$$= \prod_{i=1}^{n} \left(\frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(x_i - m)^2 \right] \right) = \tag{11.6.14}$$

$$= (2\pi\sigma^2)^{-n/2} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - m)^2 \right]$$

sau

$$L(m \mid n, x_1, \ldots, x_n, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left[-\frac{ns^2(m)}{2\sigma^2} \right]. \tag{11.6.15}$$

Din (11.6.15), ținînd seama de relația (11.6.7), obținem

$$L(m \mid n, x_1, \ldots, x_n, \sigma^2) = \tag{11.6.16}$$

$$= (2\pi\sigma^2)^{-n/2} \exp\left[-\frac{n(\bar{x} - m)^2}{2\sigma^2} \right] \exp\left[-\frac{(n-1)s^2}{2\sigma^2} \right],$$

de unde se deduce imediat relația căutată (11.6.12).

A doua afirmație a teoremei rezultă imediat din expresia (11.6.12) a funcției de verosimilitate.

Teorema 11.6.2. *În cazul cînd* σ^2 *este necunoscut, iar* m *este cunoscut, funcția de verosimilitate pentru* σ^2 *este*

$$L(\sigma^2 \mid n, x_1, \ldots, x_n, m) = k_2(\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} n s^2(m)\right), \tag{11.6.17}$$

unde

$$k_2 = (2\pi)^{-n/2}. \tag{11.6.18}$$

În plus, statistica $(n, s^2(m))$ *este suficientă pentru* σ^2.

Demonstrație. Ținînd seama de (11.6.14) putem scrie

$$L(\sigma^2 \mid n, x_1, \ldots, x_n, m) = (2\pi)^{-n/2} (\sigma^2)^{-n/2} \exp\left(-\frac{n}{2\sigma^2} s^2(m)\right),$$

de unde rezultă imediat (11.6.17), pe baza căreia obținem evident și a doua afirmație a teoremei.

Teorema 11.6.3. *În cazul cînd atît* m *cît și* σ^2 *sînt necunoscuți, funcția de verosimilitate pentru* (m, σ^2) *este*

$$L(m, \sigma^2 \mid n, x_1, \ldots, x_n) =$$

$$= k_3(\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}(n-1)s^2\right) \exp\left(-\frac{1}{2} n(\bar{x} - m)^2/\sigma^2\right), \tag{11.6.19}$$

unde

$$k_3 = (2\pi)^{-n/2}. \tag{11.6.20}$$

În plus, statistica $(n, \bar{x}, s^2)$ *este suficientă pentru* (m, σ^2).

Demonstrație. Ambele afirmații ale teoremei se deduc imediat din relația (11.6.16).

Scanned with OKEN Scanner

11.7. PROCESUL NORMAL MULTIDIMENSIONAL

11.7.1. INTRODUCERE

Vom relua problemele considerate în § 11.6 pentru cazul procesului normal multidimensional $(\tilde{\mathbf{x}}_i)$, unde $\tilde{\mathbf{x}}_i$ sînt vectori aleatori k-dimensionali independenți și identic repartizați din clasa $N(\mathbf{m}, \boldsymbol{\delta}^2)$. În cele ce urmează vom considera că vectorii aleatori $\tilde{\mathbf{x}}_i$ sînt nesingulari și deci au densitatea de probabilitate $f_N(\mathbf{x}|\mathbf{m}, \boldsymbol{\delta}^2)$.

Este clar că în ceea ce privește parametrii $\mathbf{m}$ și $\boldsymbol{\delta}^2$ putem considera următoarele cazuri posibile: $\mathbf{m}$ este cunoscut, iar $\boldsymbol{\delta}^2$ este necunoscut; $\mathbf{m}$ este necunoscut, iar $\boldsymbol{\delta}^2$ este cunoscut; $\mathbf{m}$ și $\boldsymbol{\delta}^2$ sînt necunoscuți. Este posibil de asemenea ca numai o parte dintre componentele lui $\mathbf{m}$ sau (și) o parte dintre elementele matricei $\boldsymbol{\delta}^2$ să fie cunoscute. În cele ce urmează ne vom limita la considerarea cazului cînd elementele matricei $\boldsymbol{\delta}^2$ sînt cunoscute în afara unei constante multiplicative. În acest caz definim

$$\sigma_*^2 = (\det \boldsymbol{\delta}^2)^{\frac{1}{k}} \tag{11.7.1}$$

și

$$\boldsymbol{\delta}_*^2 = (\sigma_*^2)^{-1} \boldsymbol{\delta}^2, \tag{11.7.2}$$

de unde urmează imediat că $\det \boldsymbol{\delta}_*^2 = 1$, $\boldsymbol{\delta}^2 = \sigma_*^2 \boldsymbol{\delta}_*^2$ și $\boldsymbol{\delta}_*^{-2} = \sigma_*^2 \boldsymbol{\delta}^{-2}$.

În cele ce urmează vom considera numai următoarele cazuri posibile:

Cazul A: $\mathbf{m}$ este necunoscut, iar $\boldsymbol{\delta}^2$ este cunoscut;

Cazul B: $\mathbf{m}$ este necunoscut, σ_*^2 este necunoscut, iar $\boldsymbol{\delta}_*^2$ este cunoscut;

Cazul C: $\mathbf{m}$ este cunoscut, σ_*^2 este necunoscut, iar $\boldsymbol{\delta}_*^2$ este cunoscut;

Cazul D: $\mathbf{m}$ este cunoscut, iar $\boldsymbol{\delta}^2$ este necunoscut;

Cazul E: $\mathbf{m}$ și $\boldsymbol{\delta}^2$ sînt necunoscuți.

Statisticile pe care le considerăm pentru procesul normal multidimensional sînt următoarele:

$$\tilde{\bar{\mathbf{x}}} = \frac{1}{n} \sum_{i=1}^{n} \tilde{\mathbf{x}}_i; \tag{11.7.3}$$

Scanned with OKEN Scanner

$$\tilde{s}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\tilde{x}_i - \tilde{\tilde{x}})(\tilde{x}_i - \tilde{\tilde{x}})'; \qquad (11.7.4)$$

$$\tilde{s}^2(m) = \frac{1}{n}\sum_{i=1}^{n}(\tilde{x}_i - m)(\tilde{x}_i - m)'; \qquad (11.7.5)$$

$$\tilde{s}_*^2 = \frac{1}{k(n-1)}\sum_{i=1}^{n}(\tilde{x}_i - \tilde{\tilde{x}})'\,\delta_*^{-2}(\tilde{x}_i - \tilde{\tilde{x}}); \qquad (11.7.6)$$

$$\tilde{s}_*^2(m) = \frac{1}{kn}\sum_{i=1}^{n}(\tilde{x}_i - m)'\,\delta_*^{-2}(\tilde{x}_i - m). \qquad (11.7.7)$$

Este clar că $\tilde{s}^2(m)$, $\tilde{s}_*^2$ și $\tilde{s}_*^2(m)$ sînt statistici numai pentru cazurile în care respectiv m, δ_*^2 și (m, δ_*^2) sînt cunoscute.

11.7.2. FUNCȚII DE VEROSIMILITATE

Vom prezenta mai întîi un rezultat preliminar din care se obțin apoi ușor funcțiile de verosimilitate pentru un proces $(\tilde{x}_i)$ normal de parametrii m și δ^2, în cazul cînd procesul de oprire este neinformativ asupra parametrului în discuție.

Lema 11.7.1. *Dacă* $\tilde{x}_1, \ldots, \tilde{x}_n$ *este o selecție de volum* n *din* $\tilde{x} \in N(m, \delta^2)$, *atunci densitatea de probabilitate* $f(x_1, \ldots, x_n \mid m, \delta^2)$ *este de forma*

$$f(x_1, \ldots, x_n | m, \delta^2) = (2\pi)^{-kn/2}(\det \delta^2)^{-n/2}\exp\left(-\frac{1}{2}\psi\right), \qquad (11.7.8)$$

unde ψ *este valoarea comună a expresiilor următoare*:*

$$\psi_1 = n \cdot \mathrm{ur}\,(\delta^{-2}s^2(m)); \qquad (11.7.9)$$

* Pentru orice matrice pătratică A se definește *urma* acestei matrice și se notează cu $\mathrm{ur}(A)$ suma elementelor situate pe diagonala principală. Vezi Anexa, § A.4.

$$\psi_2 = (n-1)\,\mathrm{ur}(\delta^{-2}s^2) + n(\bar{x}-m)'\,\delta^{-2}(\bar{x}-m); \qquad (11.7.10)$$

$$\psi_3 = k(n-1)\,s_*^2/\sigma_*^2 + n(\bar{x}-m)'\,\delta_*^{-2}(\bar{x}-m)/\sigma_*^2 \qquad (11.7.11)$$

şi

$$\psi_4 = kns_*^2(m)/\sigma_*^2. \qquad (11.7.12)$$

Demonstraţie. Vom arăta mai întîi că $\psi = \psi_1$. Într-adevăr, avem

$$f(x_1, \ldots, x_n | m, \delta^2) = \prod_{i=1}^{n} f_N(x_i | m, \delta^2) =$$

$$= \prod_{i=1}^{n} \frac{\exp\left[-\frac{1}{2}(x_i - m)'\,\delta^{-2}(x_i - m)\right]}{(2\pi)^{k/2}(\det \delta^2)^{1/2}} =$$

$$= \frac{\exp\left[-\frac{1}{2}\sum_{i=1}^{n} \mathrm{ur}\,(\delta^{-2}(x_i - m)(x_i - m)')\right]}{(2\pi)^{kn/2}(\det \delta^2)^{n/2}} =$$

$$= \frac{\exp\left[-\frac{1}{2}\,\mathrm{ur}\left(\delta^{-2}\sum_{i=1}^{n}(x_i - m)(x_i - m)'\right)\right]}{(2\pi)^{kn/2}(\det \delta^2)^{n/2}} =$$

$$= \frac{\exp\left[-\frac{1}{2}\,\mathrm{ur}\,(\delta^{-2}\,ns^2(m))\right]}{(2\pi)^{kn/2}(\det \delta^2)^{n/2}} = \frac{\exp\left[-\frac{1}{2}n \cdot \mathrm{ur}\,(\delta^{-2}s^2(m))\right]}{(2\pi)^{kn/2}(\det \delta^2)^{n/2}},$$

de unde decurge imediat afirmaţia căutată.

Vom arăta în al doilea rînd că $\psi_1 = \psi_2$. Într-adevăr, avem

$$\psi_1 = n \cdot \mathrm{ur}\,(\delta^{-2}s^2(m)) = \mathrm{ur}\left[\delta^{-2}\sum_{i=1}^{n}(x_i - m)(x_i - m)'\right]. \qquad (11.7.13)$$

Să observăm însă că putem scrie

$$\sum_{i=1}^{n}(\mathbf{x}_i-\mathbf{m})(\mathbf{x}_i-\mathbf{m})'=\sum_{i=1}^{n}[(\mathbf{x}_i-\bar{\mathbf{x}})+(\bar{\mathbf{x}}-\mathbf{m})][(\mathbf{x}_i-\bar{\mathbf{x}})+$$

$$+(\bar{\mathbf{x}}-\mathbf{m})]'=\sum_{i=1}^{n}(\mathbf{x}_i-\bar{\mathbf{x}})(\mathbf{x}_i-\bar{\mathbf{x}})'+\sum_{i=1}^{n}(\mathbf{x}_i-\bar{\mathbf{x}})(\bar{\mathbf{x}}-\mathbf{m})'+$$

$$+\sum_{i=1}^{n}(\bar{\mathbf{x}}-\mathbf{m})(\mathbf{x}_i-\bar{\mathbf{x}})'+\sum_{i=1}^{n}(\bar{\mathbf{x}}-\mathbf{m})(\bar{\mathbf{x}}-\mathbf{m})'=$$

$$=\sum_{i=1}^{n}(\mathbf{x}_i-\bar{\mathbf{x}})(\mathbf{x}_i-\bar{\mathbf{x}})'+n(\bar{\mathbf{x}}-\mathbf{m})(\bar{\mathbf{x}}-\mathbf{m})', \quad (11.7.14)$$

deoarece

$$\sum_{i=1}^{n}(\mathbf{x}_i-\bar{\mathbf{x}})(\bar{\mathbf{x}}-\mathbf{m})'=\left[\sum_{i=1}^{n}(\mathbf{x}_i-\bar{\mathbf{x}})\right](\bar{\mathbf{x}}-\mathbf{m})'=0$$

și, analog,

$$\sum_{i=1}^{n}(\bar{\mathbf{x}}-\mathbf{m})(\mathbf{x}_i-\bar{\mathbf{x}})'=(\bar{\mathbf{x}}-\mathbf{m})\sum_{i=1}^{n}(\mathbf{x}_i-\bar{\mathbf{x}})'=0.$$

Din relațiile (11.7.13) și (11.7.14) obținem deci

$$\psi_1=\mathrm{ur}\left[\boldsymbol{\delta}^{-2}\sum_{i=1}^{n}(\mathbf{x}_i-\bar{\mathbf{x}})(\mathbf{x}_i-\bar{\mathbf{x}})'\right]+\mathrm{ur}[\boldsymbol{\delta}^{-2}n(\bar{\mathbf{x}}-\mathbf{m})(\bar{\mathbf{x}}-\mathbf{m})']=$$

$$=\mathrm{ur}[\boldsymbol{\delta}^{-2}(n-1)\,\mathbf{s}^2]+n\cdot\mathrm{ur}\,[\boldsymbol{\delta}^{-2}(\bar{\mathbf{x}}-\mathbf{m})(\bar{\mathbf{x}}-\mathbf{m})']=$$

$$=(n-1)\,\mathrm{ur}\,[\boldsymbol{\delta}^{-2}\mathbf{s}^2]+n\cdot(\bar{\mathbf{x}}-\mathbf{m})'\,\boldsymbol{\delta}^{-2}(\bar{\mathbf{x}}-\mathbf{m})=\psi_2,$$

ceea ce dealtfel trebuia demonstrat.

Să arătăm acum că $\psi_2=\psi_3$. Într-adevăr, putem scrie

$$\psi_2=\mathrm{ur}\left[\boldsymbol{\delta}^{-2}\sum_{i=1}^{n}(\mathbf{x}_i-\bar{\mathbf{x}})(\mathbf{x}_i-\bar{\mathbf{x}})'\right]+n(\bar{\mathbf{x}}-\mathbf{m})'\,\boldsymbol{\delta}^{-2}(\bar{\mathbf{x}}-\mathbf{m})=$$

$$=\sum_{i=1}^{n}\mathrm{ur}\,[\boldsymbol{\delta}^{-2}(\mathbf{x}_i-\bar{\mathbf{x}})(\mathbf{x}_i-\bar{\mathbf{x}})']+n(\bar{\mathbf{x}}-\mathbf{m})'\,\boldsymbol{\delta}^{-2}(\bar{\mathbf{x}}-\mathbf{m})=$$

$$=\sum_{i=1}^{n}(\mathbf{x}_i-\bar{\mathbf{x}})'\,\boldsymbol{\delta}^{-2}(\mathbf{x}_i-\bar{\mathbf{x}})+n(\bar{\mathbf{x}}-\mathbf{m})'\,\boldsymbol{\delta}^{-2}(\bar{\mathbf{x}}-\mathbf{m}).$$

Scanned with OKEN Scanner

Ținînd seama de (11.7.2) putem scrie

$$\psi_2 = \frac{1}{\sigma_*^2}\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})'\,\boldsymbol{\delta}_*^{-2}(\mathbf{x}_i - \bar{\mathbf{x}}) + n(\bar{\mathbf{x}} - \mathbf{m})'\,\boldsymbol{\delta}_*^{-2}(\bar{\mathbf{x}} - \mathbf{m})/\sigma_*^2 =$$

$$= \frac{k(n-1)s_*^2}{\sigma_*^2} + n(\bar{\mathbf{x}} - \mathbf{m})'\,\boldsymbol{\delta}_*^{-2}(\bar{\mathbf{x}} - \mathbf{m})/\sigma_*^2 = \psi_3.$$

În sfîrșit, pentru a încheia demonstrația lemei, vom arăta că $\psi_1 = \psi_4$. Într-adevăr, putem scrie

$$\psi_1 = \mathrm{ur}\left[\boldsymbol{\delta}^{-2}\sum_{i=1}^{n}(\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})'\right] =$$

$$= \sum_{i=1}^{n}\mathrm{ur}[\boldsymbol{\delta}^{-2}(\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})]' = \sum_{i=1}^{n}(\mathbf{x}_i - \mathbf{m})'\,\boldsymbol{\delta}^{-2}(\mathbf{x}_i - \mathbf{m}) =$$

$$= \frac{1}{\sigma_*^2}\sum_{i=1}^{n}(\mathbf{x}_i - \mathbf{m})\,\boldsymbol{\delta}_*^{-2}(\mathbf{x}_i - \mathbf{m}) = \frac{\mathrm{k}n s_*^2(\mathbf{m})}{\sigma_*^2} = \psi_4$$

și lema este complet demonstrată.

Teorema 11.7.1. (Cazul A). *În cazul cînd* $\mathbf{m}$ *este necunoscut, iar* $\boldsymbol{\delta}^2$ *este cunoscut, funcția de verosimilitate pentru* $\mathbf{m}$ *este*

$$\mathrm{L}(\mathbf{m} \mid \mathrm{n}, \mathbf{x}_1, \ldots, \mathbf{x}_n, \boldsymbol{\delta}^2) = \mathrm{k}_1 \exp\left[-\frac{1}{2}\mathrm{n}(\bar{\mathbf{x}} - \mathbf{m})'\,\boldsymbol{\delta}^{-2}(\bar{\mathbf{x}} - \mathbf{m})\right], \tag{11.7.15}$$

unde

$$\mathrm{k}_1 = \frac{\exp\left[-\dfrac{(\mathrm{n}-1)}{2}\mathrm{ur}(\boldsymbol{\delta}^{-2}\mathbf{s}^2)\right]}{(2\pi)^{\mathrm{k}n/2}\,(\det\boldsymbol{\delta}^2)^{n/2}}. \tag{11.7.16}$$

În plus, statistica $(\mathrm{n}, \bar{\mathbf{x}})$ *este suficientă pentru* $\mathbf{m}$.

Demonstrație. Deoarece funcția de verosimilitate coincide cu $f(\mathbf{x}_1, \ldots, \mathbf{x}_n \mid \mathbf{m}, \boldsymbol{\delta}^2)$, formula (11.6.15) rezultă imediat din lema precedentă înlocuind ψ cu ψ_2. Afirmația a doua a teoremei urmează de asemenea evident din formula (11.7.15).

Scanned with OKEN Scanner

Teorema 11.7.2 (Cazul B). *În cazul cînd* $\mathbf{m}$ *este necunoscut,* σ_*^2 *este necunoscut, iar* $\boldsymbol{\delta}_*^2$ *este cunoscut, funcția de verosimilitate pentru* $(\mathbf{m}, \sigma_*^2)$ *este*

$$L(\mathbf{m}, \sigma_*^2 \mid n, \mathbf{x}_1, \ldots, \mathbf{x}_n, \boldsymbol{\delta}_*^2) =$$

$$= k_2(\sigma_*^2)^{-kn/2} \exp\left[-\frac{k(n-1)}{2} s_*^2/\sigma_*^2\right] \exp\left[-\frac{1}{2} n(\bar{\mathbf{x}} - \right.$$

$$\left. - \mathbf{m})' \boldsymbol{\delta}_*^{-2}(\bar{\mathbf{x}} - \mathbf{m})\right], \qquad (11.7.17)$$

unde

$$k_2 = (2\pi)^{-kn/2}. \qquad (11.7.18)$$

În plus, statistica $(n, \bar{\mathbf{x}}, s_*^2)$ *este suficientă pentru* $(\mathbf{m}, \sigma_*^2)$.

Demonstrație. Afirmația (11.7.17) se obține din lema 11.7.1 pentru $\psi = \psi_3$, cum se vede ușor, iar afirmația privind suficiența statisticii $(n, \bar{\mathbf{x}}, s_*^2)$ pentru $(\mathbf{m}, \sigma_*^2)$ este evidentă din forma (11.7.17) a funcției de verosimilitate.

Teorema 11.7.3. (Cazul C). *Dacă* σ_*^2 *este necunoscut, iar* $\mathbf{m}$ *și* $\boldsymbol{\delta}_*^2$ *sînt cunoscuți, funcția de verosimilitate pentru* σ_*^2 *este*

$$L(\sigma_*^2 \mid n, \mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{m}, \boldsymbol{\delta}_*^2) =$$

$$= k_3(\sigma_*^2)^{-kn/2} \exp\left[-\frac{1}{2} kn s_*^2(\mathbf{m})/\sigma_*^2\right], \qquad (11.7.19)$$

unde

$$k_3 = (2\pi)^{-kn/2}. \qquad (11.7.20)$$

În plus, statistica $(n, s_*^2(\mathbf{m}))$ *este suficientă pentru* σ_*^2.

Demonstrație. Urmează imediat din lema 11.7.1 pentru $\psi = \psi_4$.

Teorema 11.7.4. (Cazul D). *Dacă* $\boldsymbol{\delta}^2$ *este necunoscut, iar* $\mathbf{m}$ *este cunoscut, funcția de verosimilitate pentru* $\boldsymbol{\delta}^2$ *este*

$$L(\boldsymbol{\delta}^2 \mid n, \mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{m}) =$$

$$= k_4(\det \boldsymbol{\delta}^2)^{-n/2} \exp\left[-\frac{1}{2} n \cdot ur(\boldsymbol{\delta}^{-2} s^2(\mathbf{m}))\right], \qquad (11.7.21)$$

Scanned with OKEN Scanner

unde

$$k_4 = (2\pi)^{-kn/2}. \tag{11.7.22}$$

În plus, statistica $(n, s^2(m))$ *este suficientă pentru* δ^2.

Demonstrație. Urmează imediat din lema 11.7.1 pentru $\psi = \psi_1$.

Teorema 11.7.5. (Cazul E). *Dacă* m *și* δ^2 *sînt necunoscuți, atunci funcția de verosimilitate pentru* (m, δ^2) *este*

$$L(m, \delta^2 \mid n, x_1, \ldots, x_n) =$$

$$= k_5(\det \delta^2)^{-n/2} \exp\left[-\frac{(n-1)}{2} \operatorname{ur}(\delta^{-2}s^2)\right] \exp\left[-\frac{n}{2}(\bar{x} - m)' \delta^{-2}(\bar{x} - m)\right], \tag{11.7.23}$$

unde

$$k_5 = (2\pi)^{-kn/2}. \tag{11.7.24}$$

În plus, statistica $(n, \bar{x}, s^2)$ *este suficientă pentru* (m, δ^2).

Demonstrație. Rezultă imediat din lema 11.7.1 pentru $\psi = \psi_2$

Capitolul 12

REPARTIȚII CLASICE DISCRETE

12.1. INTRODUCERE

După cum am mai menționat (vezi § 10.1), studiul fenomenelor aleatoare apărute în diverse probleme practice a dus la punerea în evidență a unor repartiții care se întîlnesc mai frecvent și pe care le vom numi *repartiții clasice*.

În aplicațiile practice joacă un rol deosebit clasa variabilelor aleatoare discrete și clasa variabilelor aleatoare continue. Toate repartițiile clasice corespund unor variabile sau vectori aleatori din aceste clase. Din acest motiv, în cele ce urmează vom prezenta mai întîi repartițiile clasice discrete și apoi repartițiile clasice continue, cu excepția repartiției normale, care a fost studiată amănunțit în capitolul 10.

12.2. REPARTIȚIA DEGENERATĂ

Să considerăm o variabilă aleatoare $\tilde{x}$ pe cîmpul borelian de probabilitate $(E, \mathcal{K}, P)$ cu proprietatea că există $a \in R$ astfel încît

$$P(\tilde{x} = a) = 1.$$

Rezultă atunci că $\tilde{x}$ este o variabilă aleatoare discretă a cărei singură valoare posibilă este a. Funcția sa de masă este evident

$$f_a(x) = \begin{cases} 1 & \text{pentru } x = a, \\ 0 & \text{pentru } x \neq a, \end{cases}$$

Scanned with OKEN Scanner

iar funcția sa de repartiție este

$$F_a(x) = \begin{cases} 0 \text{ pentru } x \leqslant a, \\ 1 \text{ pentru } x > a. \end{cases}$$

Se deduce atunci ușor că media și dispersia acestei variabile aleatoare sînt respectiv

$$M(\tilde{x}) = a, \quad V(\tilde{x}) = 0,$$

iar funcția caracteristică este

$$\varphi_{\tilde{x}}(t) = e^{ita}.$$

Se poate arăta că, reciproc, dacă dispersia unei variabile aleatoare $\tilde{x}$ este nulă, atunci există $a \in R$ cu proprietatea că este singura valoare posibilă a lui $\tilde{x}$.

Repartiția pusă în evidență mai sus este numită repartiția *degenerată în* $x = a$; mulțimea variabilelor aleatoare cu această repartiție va fi notată prin H_a.

Să observăm de asemenea că funcția de repartiție F a unei variabile aleatoare discrete $\tilde{x}$ cu mulțimea valorilor $(x_i)_{i \in I}$ poate fi scrisă sub forma

$$F(x) = \sum_{i \in I} f_{\tilde{x}}(x_i) F_{x_i}(x), \tag{12.2.1}$$

unde $f_{\tilde{x}}$ este funcția de masă a lui $\tilde{x}$.

În cazul particular cînd $\tilde{x}$ ia numai valorile 0 și 1, funcția sa de repartiție se poate exprima în modul următor:

$$F(x) = f_{\tilde{x}}(0) F_0(x) + f_{\tilde{x}}(1) F_1(x) = pF_1(x) + qF_0(x),$$

unde am notat $p = f_{\tilde{x}}(1)$ și $q = f_{\tilde{x}}(0)$; evident, avem $p, q \geqslant 0$ și $p+q = 1$. Funcția de masă a acestei variabile aleatoare se poate scrie la rîndul său sub forma

$$f_{\tilde{x}}(x) = \begin{cases} p^x q^x & \text{dacă } x \in \{0, 1\}, \\ 0 & \text{dacă } x \notin \{0, 1\}. \end{cases} \tag{12.2.2}$$

Funcția caracteristică a acestei variabile aleatoare este

$$\varphi_{\tilde{x}}(t) = p e^{it} + q. \tag{12.2.3}$$

Scanned with OKEN Scanner

În sfîrșit, valoarea medie și dispersia acestei variabile aleatoare sînt respectiv

$$M(\tilde{x}) = 1 \cdot p + 0 \cdot q = p, \tag{12.2.4}$$

și

$$V(\tilde{x}) = M((\tilde{x}-p)^2) = p(1-p)^2+q(0-p)^2 = pq. \tag{12.2.5}$$

12.3. PROCESUL BERNOULLI ȘI REPARTIȚII ASOCIATE

O primă serie de repartiții discrete apare în legătură cu o schemă de experiment statistic imaginată de J. Bernoulli (1654—1705), care este, de exemplu, realizată de o succesiune de aruncări ale unei monede sau de extrageri *independente* dintr-o urnă care conține bile de două culori (independența implicînd necesitatea ca, după fiecare extragere, bila extrasă să fie reintrodusă în urnă înainte de a se trece la extragerea următoare). Schema lui Bernoulli, numită și *schema bilei revenite (repuse)*, a avut o mare importanță în aplicațiile practice ale teoriei probabilităților.

În general, vom numi *proces Bernoulli* un proces generator de date care conduce la variabile aleatoare $\tilde{x}_1, \tilde{x}_2, \ldots$ independente și identic repartizate care iau numai valorile 1 și 0 cu probabilitățile p și $1-p$, unde $p \in (0, 1)$ este numit *parametrul procesului*.

Este util să observăm că funcția de masă a unei variabile aleatoare $\tilde{x}_i$ din acest proces poate fi scrisă sub forma

$$f_{\tilde{x}_i}(x|p) = \begin{cases} p^x(1-p)^x & \text{dacă } x \in \{0, 1\}, \\ 0 & \text{dacă } x \notin \{0, 1\}, \end{cases} \tag{12.3.1}$$

conform formulei (12.2.2).

Procesul Bernoulli apare, în general, în experimente în care la fiecare observație (probă) poate să apară un eveniment A sau contrarul său A^c, valoarea lui $\tilde{x}_i$ fiind luată egală cu 1 sau zero după cum în observația de rang i apare A sau A^c, probabilitatea de apariție a lui A într-o probă fiind p. În exemplul cu moneda, $\tilde{x}_i$ ia valoarea 1 sau 0 după cum în aruncarea de rang i apare stema sau valoarea, probabilitatea de apariție

a stemei într-o aruncare fiind p. Analog, în exemplul cu urna, care conține a bile albe și r bile roșii, $\tilde{x}_i$ este egal cu 1 sau cu 0 după cum în extragerea de rang i din urnă apare o bilă albă sau o bilă roșie, probabilitatea de apariție a unei bile albe fiind $p = a/(a+r)$.

Acest proces este utilizat frecvent în probleme de controlul calității loturilor de produse care conțin o proporție p de unități defecte, în probleme de sondaje privind preferințele unei anumite populații pentru un produs dat, p fiind proporția celor care preferă acest produs etc.

Dacă se consideră o selecție de volum n dintr-un proces Bernoulli de parametru $p \in (0, 1)$, atunci funcția de masă a vectorului aleator $(\tilde{x}_1, ..., \tilde{x}_n)$ se poate scrie sub forma

$$f(x_1, ..., x_n \mid n, p) = f(x_1 \mid p) \dots f(x_n \mid p)$$

și deci, ținînd seama de (12.3.1), obținem

$$f(x_1, ..., x_n \mid n, p) = \begin{cases} p^r(1-p)^r & \text{dacă } r \in \{0, 1, ..., n\}, \\ 0 & \text{dacă } r \notin \{0, 1, ..., n\}, \end{cases} \tag{12.3.2}$$

unde $r = x_1 + ... + x_n$.

12.3.1. REPARTIȚIA BINOMIALĂ

Să considerăm un proces de oprire neinformativ asupra parametrului p care constă în fixarea în prealabil a numărului n de observații ce urmează a se efectua. Este clar că variabila aleatoare

$$\tilde{x}_{(n)} = \tilde{x}_1 + ... + \tilde{x}_n \tag{12.3.3}$$

reprezintă numărul de apariții ale evenimentului A în cele n observații ce urmează a se efectua. Variabila aleatoare $\tilde{x}_{(n)}$ este tot o variabilă aleatoare discretă, avînd mulțimea valorilor $0, 1, ..., n$, funcția sa de masă fiind de forma

$$f_{\tilde{x}_{(n)}}(r) = P(\tilde{x}_{(n)} = r) = \sum f(x_1, ..., x_n \mid n, p), \tag{12.3.4}$$

unde suma se face în raport cu toți vectorii $(x_1, ..., x_n)'$ în care exact r, $0 \leqslant r \leqslant n$, componente sînt egale cu 1 (celelalte

fiind zero). Cum în această sumă sînt evident C_n^r termeni, din (12.3.4) și (12.3.2) rezultă că

$$f_{\tilde{x}_{(n)}}(r)=\begin{cases} C_n^r p^r (1-p)^{n-r} & \text{dacă } r\in\{0, 1, \ldots, n\}, \\ 0 & \text{dacă } r\notin\{0, 1, \ldots, n\}. \end{cases} \tag{12.3.5}$$

Funcția de masă (12.3.5) este numită *funcția de masă binomială* și va fi notată prin $f_b(r\,|\,n, p)$, unde $n\in N$ și $p\in(0, 1)$ sînt *parametrii repartiției binomiale*. Mulțimea variabilelor aleatoare care au o funcție binomială de masă de parametrii n și p va fi notată prin $B(n, p)$.

Din modul în care a fost pusă în evidență această repartiție rezultă că suma a n variabile aleatoare $\tilde{x}_1, \ldots, \tilde{x}_n$ provenind dintr-un proces Bernoulli de parametru p este o variabilă aleatoare binomială de parametrii n și p.

Funcția de repartiție a unei variabile $\tilde{x}\in B(n, p)$ este

$$F_b(x\,|\,n, p) = \sum_{r=0}^{[x]-1} C_n^r p^r (1-p)^{n-r}, \tag{12.3.6}$$

unde $[x]$ este cel mai mare întreg mai mic sau egal cu x.

Funcția caracteristică a unei variabile aleatoare $\tilde{x}\in B(n, p)$ este, evident,

$$\varphi_b(t|n, p) = (p\,e^{it} + q)^n, \tag{12.3.7}$$

unde $q = 1-p$, ținînd seama că $\tilde{x}$ poate fi reprezentată ca suma a n variabile aleatoare independente dintr-un proces Bernoulli de parametru p.

După cum se vede ușor, momentele și momentele centrate pînă la ordinul al doilea pentru o variabilă aleatoare $\tilde{x}\in B(n, p)$ sînt:

$$M(\tilde{x}) = np;\ M(\tilde{x}^2) = n^2p^2 + npq;\ V(\tilde{x}) = npq. \tag{12.3.8}$$

Valoarea cea mai probabilă (modul) a repartiției binomiale este numărul întreg λ care îndeplinește condițiile

$$np - q \leqslant \lambda \leqslant np + p. \tag{12.3.9}$$

Scanned with OKEN Scanner

Într-adevăr, din

$$\frac{f_b(r \mid n, p)}{f_b(r-1 \mid n; p)} = \frac{n-r+1}{r} \cdot \frac{p}{q}$$

și

$$\frac{f_b(r+1 \mid n, p)}{f_b(r \mid n, p)} = \frac{n-r}{r+1} \cdot \frac{p}{q}$$

pentru $0 \leqslant r \leqslant n$, rezultă că λ trebuie să satisfacă inegalitățile:

$$\frac{n-\lambda+1}{\lambda} \cdot \frac{p}{q} \geqslant 1; \quad \frac{n-\lambda}{\lambda+1} \cdot \frac{p}{q} \leqslant 1,$$

de unde se obține imediat (12.3.9).

În general numărul $np-q$ nu este întreg, de unde rezultă că nici $np+p$ nu este întreg (altfel, diferența $np+p-(np-q)= =p+q=1$ ar fi un număr întreg, ceea ce ar contrazice faptul că $np-q$ nu este întreg).

În cazul particular cînd $np-q$ este întreg, rezultă că $np+p$ este de asemenea întreg și deci repartiția binomială corespunzătoare este bimodală.

Pentru valorile mari ale lui n, caz întîlnit frecvent în aplicații, calculul probabilităților $f_b(r|n, p)$ este destul de dificil. Se poate obține o valoare aproximativă a acestor probabilități dacă utilizăm *formula lui Stirling*

$$r! = (2\pi r)^{1/2} \cdot r^{r/2} \, e^{-r+\theta(r)}, \tag{12.3.10}$$

unde $\theta(r)$ verifică inegalitatea

$$|\theta(r)| < 1/12r. \tag{12.3.11}$$

Într-adevăr, utilizînd această formulă, rezultă că

$$f_b(r|n, p) = \frac{n!}{r!(n-r)!} p^r q^{n-r} \approx$$

$$\approx \sqrt{\frac{n}{2\pi r(n-r)}} \cdot \frac{n^n p^r q^{n-r}}{r^r (n-r)^{n-r}} e^{\theta},$$

unde $\theta = \theta(n) - \theta(r) - \theta(n-r)$ verifică inegalitatea

$$|\theta| < \frac{1}{12}\left(\frac{1}{n} + \frac{1}{r} + \frac{1}{n-r}\right),$$

conform relației (12.3.11). Un exercițiu simplu ne conduce la formula

$$f_b(r \mid n, p) \approx (2\pi npq)^{1/2} e^{\frac{(r-np)^2}{-2npq}}, \qquad (12.3.12)$$

unde $r \in \{0, 1, \ldots, n\}$.

Observînd că np și npq sînt respectiv media m și dispersia σ^2 ale unei variabile aleatoare binomiale $\tilde{x} \in B(n, p)$, din (12.3.12) rezultă că $f_b(r \mid n, p)$ se poate aproxima cu valoarea densității de probabilitate normale cu parametrii $m = np$ și $\sigma^2 = npq$, adică $f_N(r \mid np, npq)$. Prin urmare

$$f_b(r \mid n, p) \approx f_N(r \mid np, npq)$$

pentru $r \in \{0, 1, \ldots, n\}$. Rezultatul (12.3.12) este deci o teoremă limită locală și se putea obține imediat din teorema 12.6.3 a lui Gnedenko privind repartițiile în rețea.

Se poate obține de asemenea o teoremă limită centrală pentru procesul Bernoulli. Într-adevăr, dacă $(\tilde{x}_n)$ este un proces Bernoulli de parametru $p \in (0, 1)$, atunci, din teorema 10.5.3 (vezi condiția (c) a teoremei), rezultă că variabilele aleatoare

$$\frac{\tilde{x}_{(n)} - np}{\sqrt{npq}}$$

converg în repartiție către o variabilă aleatoare normală standard. Cu alte cuvinte, o variabilă aleatoare $\tilde{x} \in B(n, p)$ este asimptotic normală, de parametrii np și npq. Rezultă atunci că, pentru n suficient de mare, putem scrie

$$P(a\sqrt{npq} < x - np < b\sqrt{npq}) \approx (2\pi)^{-1/2} \int_a^b e^{-x^2/2}\, dx. \qquad (12.3.13)$$

Folosirea în practică a acestei formule nu este totdeauna indicată. Pentru lămurirea situațiilor în care aproximarea dată de această formulă este acceptabilă dăm, fără demonstrație, un rezultat datorat lui S. N. Bernstein.

Scanned with OKEN Scanner

Teorema 12.3.1. *Dacă*

$$0 < a < (16\,npq)^{1/6}, \quad npq \geqslant 62{,}5, \tag{12.3.14}$$

atunci

$$P\left(-a + \frac{q-p}{6\sqrt{npq}}a^2 < \frac{\tilde{x} - np}{\sqrt{npq}} < a + \frac{q-p}{6\sqrt{npq}}a^2\right) =$$

$$= \frac{2}{\sqrt{2\pi}}\int_0^a e^{-x^2/2}dx + R_n,$$

unde

$$|R_n| < 2e^{-\sqrt{2npq}}.$$

Să observăm că prima condiție din (12.3.14) este îndeplinită aproape totdeauna, pentru valori mari ale lui n, deoarece în practică ne interesează numai valorile lui a mai mici decît 1. În ceea ce privește a doua condiție, este de remarcat că, pentru p și q de valori apropiate, n nu trebuie să fie prea mare. Dacă însă p și q sînt aproape de zero sau de 1, numărul n al observațiilor trebuie să fie foarte mare pentru ca (12.3.14) să fie îndeplinite.

12.3.2. REPARTIȚIA PASCAL

Să considerăm un proces Bernoulli $(\tilde{x}_n)$ de parametru $p \in (0, 1)$. În cazul repartiției binomiale, am fixat n, adică am considerat un proces de oprire (neinformativ asupra lui p) de forma (vezi § 11.3):

$$\pi(i|x_1, \ldots, x_i, p) = \begin{cases} 0 & \text{pentru } 0 \leqslant i \leqslant n-1, \\ 1 & \text{pentru } i = n. \end{cases} \tag{12.3.15}$$

Vom considera acum un al doilea proces de oprire. Mai precis, vom continua să observăm valorile variabilelor $\tilde{x}_1, \tilde{x}_2, \ldots$

pînă cînd exact r valori observate sînt egale cu 1. Acest proces de oprire este deci dat de probabilitățile de oprire:

$$\pi(i|x_1, \ldots, x_i, p) = \begin{cases} 0 \text{ pentru } \sum_{j=1}^{i} x_j < r, \\ \\ 1 \text{ pentru } \sum_{j=1}^{i} x_j \geqslant r, \end{cases} \tag{12.3.16}$$

și deci este un proces de oprire neinformativ pentru p. Experimentele statistice în care se utilizează acest tip de proces de oprire sînt numite și procedee de *selecție inversă*.

Dacă în rezultatul $(x_1, \ldots, x_n)$ al experimentului statistic avem exact r valori egale cu 1, rezultă că $x_n = 1$ (altfel, nu ne-am fi oprit la observația n) și deci orice rezultat de forma $(x_1, \ldots, x_{n-1}, 1)$ va conține exact $r-1$ valori egale cu 1 printre valorile $x_1, \ldots, x_{n-1}$, probabilitatea unui astfel de rezultat fiind evident $p^r(1-p)^{n-r}$. Cum numărul rezultatelor de forma $(x_1, \ldots, x_{n-1}, 1)$ care au exact r valori egale cu 1 este evident C_{n-1}^{r-1}, rezultă că putem scrie

$$P(\tilde{n} = n) = \begin{cases} 0 & \text{dacă } n \notin \{r, r+1, \ldots\}, \\ C_{n-1}^{r-1} p^r (1-p)^{n-r} & \text{dacă } n \in \{r, r+1, \ldots\}. \end{cases}$$

Prin urmare, numărul $\tilde{n}$ de observații necesar pentru a obține r valori egale cu 1 dintr-un proces Bernoulli de parametru $p \in (0, 1)$ este o variabilă aleatoare discretă care ia valorile $r, r+1, \ldots$ cu probabilitățile $C_{n-1}^{r-1} p^r (1-p)^{n-r}$.

Spunem că o variabilă aleatoare discretă are o *repartiție Pascal de parametrii* r *și* $p \in (0, 1)$, dacă funcția sa de masă, pe care o notăm prin $f_{Pa}(n \mid r, p)$, are forma

$$f_{Pa}(n \mid r, p) = \begin{cases} C_{n-1}^{r-1} p^r (1-p)^{n-r} & \text{pentru } n \in \{r, r+1, \ldots\}, \\ 0 & \text{pentru } n \notin \{r, r+1, \ldots\}. \end{cases} \tag{12.3.17}$$

Mulțimea variabilelor aleatoare care au repartiție Pascal de parametrii r și p va fi notată prin $Pa(r, p)$.

Din cele de mai sus rezultă că numărul de observații necesar pentru a obține r valori egale cu 1 dintr-un proces Bernoulli de parametru p este o variabilă aleatoare repartizată Pascal cu parametrii r și p.

Scanned with OKEN Scanner

Deoarece $C_{n-1}^{r-1} p^r(1-p)^{n-r}$ reprezintă termenul general al dezvoltării binomului,

$$\left(\frac{1}{p} - \frac{1-p}{p}\right)^{-r} = p^r(1-(1-p))^{-r}$$

cu exponent negativ, repartiția Pascal este numită uneori *repartiția binomială cu exponent negativ*. În cele ce urmează vom opta pentru prima denumire, rezervînd pe cea de a doua pentru o altă repartiție, care va fi pusă în evidență în § 12.2.3.

Funcția caracteristică a unei variabile aleatoare cu repartiție Pascal de parametrii r și $p \in (0, 1)$ este, conform definiției,

$$\varphi_{Pa}(t|r, p) = \sum_{n=r}^{\infty} C_{n-1}^{r-1}(p\, e^{it})\,(1-(1-p)e^{it})^{n-r} = \\ = (pe^{it})^r(1-(1-p)e^{it})^{-r}, \tag{12.3.18}$$

ținînd seama de dezvoltarea binomului $(1-(1-p)e^{it})^{-r}$ cu exponent negativ.

Se poate vedea ușor că dacă $\tilde{x}_1$ și $\tilde{x}_2$ sînt variabile aleatoare independente, unde $\tilde{x}_1 \in Pa(r_1, p)$ și $x_2 \in Pa(r_2, p)$, atunci $\tilde{x}_1 + \tilde{x}_2 \in Pa(r_1 + r_2, p)$. Într-adevăr, avem

$$\varphi_{\tilde{x}_1+\tilde{x}_2}(t) = \varphi_{Pa}(t \mid r_1, p)\varphi_{Pa}(t \mid r_2, p) = \\ = (pe^{it})^{r_1+r_2}(1-(1-p)e^{it})^{-(r_1+r_2)} = \varphi_{Pa}(t \mid r_1 + r_2, p).$$

Media și dispersia unei variabile aleatoare $\tilde{x} \in Pa(r, p)$ se deduc ușor cu ajutorul funcției caracteristice. Într-adevăr,

$$M(\tilde{x}) = (1/i)\,\varphi'_{Pa}(0 \mid r, p) = r/p, \\ M(\tilde{x}^2) = r^2/p^2 + r(1-p)/p^2 \tag{12.3.19}$$

și deci

$$V(\tilde{x}) = r(1-p)/p^2. \tag{12.3.20}$$

Repartiția limită pentru $r \to +\infty$ a unei variabile aleatoare repartizată Pascal cu parametrii r și p se poate obține cu ajutorul funcției caracteristice a variabilei reduse

$$\tilde{y} = \frac{\tilde{x} - M(\tilde{x})}{V(\tilde{x})} = \frac{p\tilde{x} - r}{\sqrt{r(1-p)}} \tag{12.3.21}$$

Scanned with OKEN Scanner

pentru care funcția caracteristică este

$$\varphi_{\tilde{y}}(t) = e^{-it\sqrt{r/(1-p)}}\, \varphi_{Pa}\left(\frac{pt}{\sqrt{r(1-p)}} \mid r,\ p\right),$$

adică

$$\varphi_{\tilde{y}}(t) = \left(\frac{1}{p}\, e^{-it\sqrt{r/(1-p)}} - \frac{1-p}{p}\, e^{it\sqrt{1/r(1-p)}}\right)^{-r}.$$

Dezvoltînd în serie pe e^{ix} și oprindu-ne la termenul de gradul al treilea putem scrie

$$e^{ix} = 1 + ix/1! + (ix)^2/2! + \theta x^3/3!,$$

unde $|\theta|<1$, obținem după calcule simple

$$\varphi_{\tilde{y}}(t) = (1 + t^2/2r - \theta t^3(2-p)/(6r^{3/2}(1-p)^{1/2}))^{-r}.$$

Trecînd la limită pentru $r\to+\infty$ se obține

$$\lim_r \varphi_{\tilde{y}}(t) = e^{-t^2/2},$$

ceea ce arată că repartiția variabilei aleatoare $\tilde{y}$ dată de (12.3.21) este asimptotic repartiția normală standard.

12.3.3. REPARTIȚIA BINOMIALĂ NEGATIVĂ

O repartiție strîns legată de repartiția Pascal de parametrii r și $p\in(0, 1)$ este repartiția numărului de zerouri care apar înaintea celui de al r-lea 1, adică repartiția variabilei aleatoare discrete $\tilde{z} = \tilde{n} - r$, unde $\tilde{n}\in Pa(r, p)$. Deoarece

$$P(\tilde{z} = z) = P(\tilde{n} = z + r),$$

rezultă că funcția de masă a variabilei $\tilde{z}$ este

$$f(z) = \begin{cases} C_{z+r-1}^{r-1}\, p^r(1-p)^z & \text{pentru } z\in\{0, 1, \ldots\}, \\ 0 & \text{pentru } z\notin\{0, 1, \ldots\}. \end{cases}$$

Spunem că o variabilă aleatoare discretă are o *repartiție binomială negativă* de parametrii r și p, dacă funcția sa de masă, pe care o vom nota prin $f_{bn}(z \mid r, p)$, este dată de relația

$$f_{bn}(z \mid r, p) = \begin{cases} C_{z+r-1}^{r-1} p^r (1-p)^z & \text{dacă } z \in \{0, 1, \ldots\}, \\ 0 & \text{dacă } z \notin \{0, 1, \ldots\}. \end{cases} \quad (12.3.22)$$

Mulțimea variabilelor aleatoare care au repartiție binomială negativă de parametrii r și p va fi notată prin $Bn(r, p)$.

Din cele de mai sus rezultă că numărul de valori egale cu zero care apar înaintea celui de al r-lea 1 într-un proces Bernoulli de parametru $p \in (0, 1)$ este o variabilă aleatoare cu repartiție binomială negativă de parametrii r și p.

Repartiția binomială negativă de parametrii $r = 1$ și p este numită și *repartiția geometrică de parametru* p, funcția sa de masă fiind notată prin $f_g(z \mid p)$, adică:

$$f_g(z \mid p) = \begin{cases} p(1-p)^z & \text{pentru } z \in \{0, 1, \ldots\}, \\ 0 & \text{pentru } z \notin \{0, 1, \ldots\}. \end{cases} \quad (12.3.23)$$

Funcția de repartiție corespunzătoare acestei repartiții este nulă pentru $x \leqslant 0$ și are valoarea

$$F_g(x \mid p) = 1 - (1-p)^{[x]+1}, \quad (12.3.24)$$

pentru $x > 0$, unde $[x]$ este cel mai mare întreg mai mic sau egal cu x.

12.3.4. *Exemple.* Vom da cîteva exemple tipice de utilizare a repartițiilor binomială, Pascal și binomială negativă.

Exemplul 1. Să considerăm un lot format din 1000 piese, dintre care 200 sînt defecte. Extragem la întîmplare, cu repunere, $n = 4$ piese. Probabilitatea ca toate piesele extrase să fie defecte (adică $r = 4$) este

$$f_b(4 \mid 4, 1/5) = C_4^4 (1/5)^4 (4/5)^0 = 1/5^4 = 0{,}0016,$$

unde $p = 200/1000 = 1/5$ este proporția pieselor defecte. Evident, această probabilitate este foarte mică și deci apariția a patru piese defecte este foarte improbabilă pentru acest lot.

Exemplul 2. Să presupunem că nu cunoaștem proporția p de piese defecte dintr-un lot de 1000 piese și că cineva ne spune că $p = 1/5$. Pentru a verifica această afirmație, extragem

Scanned with OKEN Scanner

$n = 4$ piese din lot și obținem $r = 3$ piese defecte. Problema este: acceptăm că $p = 1/5$ sau nu?

Probabilitatea ca să apară $r = 3$ piese defecte în ipoteza că $p = 1/5$ este

$$f_b(3 \mid 4,\ 1/5) = C_4^3\,(1/5)^3(4/5)^{4-3} = 0{,}0256.$$

Această probabilitate este destul de mică, dîndu-ne motive să ne îndoim de informația că $p = 1/5$. Pentru o clarificare a situației, extragem cu repunere încă 6 piese, dintre care 5 piese se dovedesc defecte. Rezultă atunci că din $n = 4+6=10$ piese extrase din lot s-au obținut $r = 3+5=8$ piese defecte. Probabilitatea acestui eveniment este

$$f_b(8 \mid 10,1/5) = C_{10}^8(1/5)^8(4/5)^2 \approx 0{,}00007$$

și deci avem motive foarte serioase să ne îndoim de faptul că $p = 1/5$.

Asupra problemei estimării parametrului p al unui proces Bernoulli vom avea ocazia să mai revenim în cele ce urmează.

Exemplul 3. Să presupunem că proporția suporterilor unui anumit candidat în alegeri este $p = 0{,}9$. Numărul n al persoanelor care trebuie chestionate pentru a găsi $r = 2$ suporteri ai acestui candidat este o variabilă aleatoare repartizată Pascal cu parametrii $r = 2$ și $p = 0{,}9$. Rezultă atunci că probabilitățile de a chestiona 2, 3, 4, 5 persoane pentru a găsi doi suporteri sînt respectiv

$$f_{Pa}(2 \mid 2;\ 0{,}9) = C_{2-1}^{2-1}(0{,}9)^2(0{,}1)^{2-2} = 0{,}81,$$

$$f_{Pa}(3 \mid 2;\ 0{,}9) = C_{3-1}^{2-1}(0{,}9)^2(0{,}1)^{3-2} = 0{,}162,$$

$$f_{Pa}(4 \mid 2;\ 0{,}9) = C_{4-1}^{2-1}(0{,}9)^2(0{,}1)^{4-2} = 0{,}0243,$$

$$f_{Pa}(5 \mid 2;\ 0{,}9) = C_{5-1}^{2-1}(0{,}9)^2(0{,}1)^{5-2} = 0{,}00324.$$

Rezultă atunci că probabilitatea ca numărul persoanelor chestionate pentru a găsi doi suporteri să fie cuprins între 2 și 5 este

$$0{,}81 + 0{,}162 + 0{,}0243 + 0{,}00324 = 0{,}99954$$

Scanned with OKEN Scanner

și deci probabilitatea de a chestiona mai mult de șase persoane pentru a găsi doi suporteri este

$$1 - 0{,}99954 = 0{,}00046.$$

Cu alte cuvinte, este foarte improbabil (46 cazuri din 10 000) ca să chestionăm mai mult de șase persoane pentru a găsi doi suporteri.

Exemplul 4. Să presupunem că 30% din gospodine preferă detergentul *Cristal*. Se organizează o anchetă privind preferințele pentru acest detergent. Numărul gospodinelor care răspund negativ înainte de a se obține $r = 10$ răspunsuri pozitive privind preferințele pentru acest detergent este o variabilă aleatoare care urmează o repartiție binomială negativă de parametrii $r = 10$ și $p = 0{,}3$.

Probabilitatea ca toate gospodinele să răspundă afirmativ (adică numărul z al răspunsurilor negative să fie nul) este

$$f_{bn}(0 \mid 10; 0{,}3) = C_{0+10-1}^{10-1}(0{,}3)^{10}\,(0{,}7)^0 \approx 6 \cdot 10^{-6}.$$

Probabilitățile ca să obținem 1, 2, 3, 4 răspunsuri negative sînt respectiv:

$$f_{bn}(1 \mid 10;\ 0{,}3) = C_{10}^{9}\,(0{,}3)^{10}(0{,}7)^1 \approx 4 \cdot 10^{-5};$$

$$f_{bn}(2 \mid 10;\ 0{,}3) = C_{11}^{9}(0{,}3)^{10}(0{,}7)^2 \approx 16 \cdot 10^{-5};$$

$$f_{bn}(3 \mid 10;\ 0{,}3) = C_{12}^{9}(0{,}3)^{10}(0{,}7)^3 \approx 45 \cdot 10^{-5};$$

$$f_{bn}(4 \mid 10;\ 0{,}3) = C_{13}^{9}(0{,}3)^{10}(0{,}7)^4 \approx 94 \cdot 10^{-5}.$$

12.4. PROCESUL BERNOULLI MULTIDIMENSIONAL

În cele ce urmează vom considera un șir (finit sau infinit) de vectori aleatori k—dimensionali $\tilde{\mathbf{x}}_1$, $\tilde{\mathbf{x}}_2$, ... independenți și identic repartizați, avînd drept valori posibile vectorii $\mathbf{e}_0 = (0, \ldots, 0)' = \mathbf{0}$, $\mathbf{e}_1 = (1, 0, \ldots, 0)'$, ..., $\mathbf{e}_k = (0, \ldots, 0, 1)'$ din R^k.

Vom nota

$$p_j = P(\tilde{\mathbf{x}}_n = \mathbf{e}_j),\ 0 \leqslant j \leqslant k, \qquad (12.4.1)$$

pentru orice $n \in N$. Evident, avem satisfăcute relațiile:

$$p_j > 0, \quad 0 \leqslant j \leqslant k;$$
$$\sum_{j=0}^{k} p_j = 1, \qquad 0 < \sum_{j=1}^{k} p_j < 1. \tag{12.4.2}$$

Vom nota prin $\mathbf{p}$ vectorul cu componentele $p_1, \ldots, p_k$.

Spunem că șirul de vectori aleatori $\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \ldots$ definit mai înainte formează un proces Bernoulli k—dimensional de parametru $\mathbf{p}$. Este evident că procesul considerat mai sus este o generalizare a procesului Bernoulli unidimensional, care se obține evident din definiția dată pentru cazul particular $k = 1$.

O interpretare imediată a acestui proces este următoarea. Să considerăm o urnă care conține bile de $k+1$ culori, notate prin $c_0, c_1, \ldots, c_k$, în proporțiile $p_0, \ldots, p_k$; evident, $p_0, p_1, \ldots, p_k$ verifică condițiile (12.4.2). Dacă facem extrageri succesive cu repunere din această urnă și introducem vectorul aleator $\tilde{\mathbf{x}}_n$ definit de

$$\tilde{\mathbf{x}}_n = \begin{cases} \mathbf{e}_0 & \text{dacă apare bilă de culoarea } c_0, \\ \mathbf{e}_1 & \text{dacă apare bilă de culoarea } c_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \mathbf{e}_k & \text{dacă apare bilă de culoarea } c_k, \end{cases}$$

pentru extragerea de rang n, atunci este clar că $\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \ldots$ formează un proces Bernoulli k—dimensional de parametru $\mathbf{p} = (p_1, \ldots, p_k)'$.

În general, dacă $(\tilde{\mathbf{x}}_n)$ este un proces Bernoulli k—dimensional de parametru $\mathbf{p} = (p_1, \ldots, p_k)'$, atunci funcția de masă a lui $\tilde{\mathbf{x}}_n$ se poate scrie sub forma

$$f(\mathbf{x}|\mathbf{p}) = \begin{cases} p_0^{x_0} p_1^{x_1} \ldots p_k^{x_k} & \text{dacă } \mathbf{x} = (x_1, \ldots, x_k)' \text{ este o valoare} \\ & \text{posibilă a vectorului aleator } \tilde{\mathbf{x}}_n, \\ 0 & \text{în caz contrar,} \end{cases} \tag{12.4.3}$$

unde

$$x_0 = 1 - (x_1 + \ldots + x_k),$$
$$p_0 = 1 - (p_1 + \ldots + p_k). \tag{12.4.4}$$

Este de asemenea clar că funcția de masă a vectorului aleator $(\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_n)'$ poate fi scrisă sub forma

$$f(\mathbf{x}_1, \ldots, \mathbf{x}_n \mid n, \mathbf{p}) = \begin{cases} p_0^{r_0} p_1^{r_1} \ldots p_k^{r_k} & \text{pentru toți } r_j \in \{0, 1, \ldots, n\}, \\ & \text{pentru care } r_0 + r_1 + \ldots + r_k = n \\ 0 & \text{în caz contrar,} \end{cases} \tag{12.4.5}$$

unde

$$r_j = \sum_{i=1}^{n} x_{ji}, \tag{12.4.6}$$

iar x_{ji} este componenta de rang j a lui $\mathbf{x}_i$, $1 \leqslant i \leqslant n$.

Este clar că în exemplul de mai sus, r_j reprezintă numărul bilelor de culoarea c_j care au apărut în urma celor n extrageri efectuate.

Vom nota prin $\mathbf{r}$ vectorul cu componentele $r_1, \ldots, r_k$ definite de (12.4.6). Cu această notație, rezultă că (12.4.6) poate fi scrisă sub forma echivalentă

$$\mathbf{r} = \sum_{i=1}^{n} \mathbf{x}_i. \tag{12.4.7}$$

Din (12.4.5) și (12.4.7) se vede ușor că $(n, \mathbf{r})$ este o statistică suficientă pentru $\mathbf{p}$.

Ca și în cazul unidimensional vom considera mai întîi un proces de oprire neinformativ asupra lui $\mathbf{p}$ caracterizat de faptul că numărul n al probelor (observațiilor) este predeterminat.

Funcția de masă a vectorului aleator

$$\tilde{\mathbf{x}}_{(n)} = \tilde{\mathbf{x}}_1 + \ldots \tilde{\mathbf{x}}_n \tag{12.4.8}$$

se obține imediat și este dată de relația

$$f_{\tilde{\mathbf{x}}_{(n)}}(\mathbf{r}) = P(\tilde{\mathbf{x}}_{(n)} = \mathbf{r}) = P(\tilde{\mathbf{x}}_1 + \ldots + \tilde{\mathbf{x}}_n = \mathbf{r}) =$$

$$= \sum f(\mathbf{x}_1, \ldots, \mathbf{x}_n \mid n, \mathbf{p}),$$

unde suma se face după toate valorile $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ care verifică condiția (12.4.7). Deoarece $f(\mathbf{x}_1, \ldots, \mathbf{x}_n \mid n, \mathbf{p})$ este constant pe mulțimea tuturor acelor $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ care verifică condiția (12.4.7), iar această mulțime are un număr de elemente egal cu

$$\frac{n!}{r_0!r_1!\ldots r_k!} \tag{12.4.9}$$

rezultă că

$$f_{\tilde{\mathbf{x}}_{(n)}}(\mathbf{r}) = \begin{cases} \dfrac{n!}{r_0!r_1!\ldots r_k!} p_0^{r_0} p_1^{r_1} \ldots p_k^{r_k} & \text{pentru toți } r_j \in \{0, 1, \ldots, n\} \\ & \text{pentru care } r_0 + \ldots + r_k = n, \\ 0 & \text{în caz contrar.} \end{cases} \tag{12.4.10}$$

Se pune în acest mod în evidență o repartiție pe care o vom numi *repartiție multinomială*. Astfel, vom spune că un vector aleator k—dimensional are o *repartiție multinomială de parametrii* n și $\mathbf{p}$ dacă funcția sa de masă este de forma (12.4.10); vom nota această funcție de masă prin $f_{Mu}(\mathbf{r} \mid n, \mathbf{p})$, iar mulțimea vectorilor aleatori care au această funcție de masă va fi notată prin $Mu(n, \mathbf{p})$. Denumirea acestei repartiții provine din faptul că valorile diferite de zero ale funcției de masă $f_{Mu}(\mathbf{r} \mid n, \mathbf{p})$ sînt chiar termenii dezvoltării polinomului $(p_0 + p_1 + \ldots + p_k)^n$.

Din cele de mai sus rezultă că suma a n vectori aleatori dintr-un proces Bernoulli k—dimensional de parametru $\mathbf{p}$ este repartizată multinomial cu parametrii n și $\mathbf{p}$.

Se poate vedea ușor că funcția caracteristică a unui vector aleator $\tilde{\mathbf{x}} \in Mu(n, \mathbf{p})$ este

$$\varphi_{Mu}(\mathbf{t} \mid \mathrm{n}, \mathbf{p}) = (\mathrm{p}_0 + \mathrm{p}_1 \mathrm{e}^{\mathrm{i}t_1} + \ldots + p_k \mathrm{e}^{\mathrm{i}t_k})^n. \tag{12.4.11}$$

Dacă vectorii aleatori independenți k — dimensionali $\tilde{\mathbf{x}}_i \in Mu(n_i, \mathbf{p})$, $1 \leqslant i \leqslant s$, atunci suma lor

$$\tilde{\mathbf{x}} = \tilde{\mathbf{x}}_1 + \ldots + \tilde{\mathbf{x}}_s$$

Scanned with OKEN Scanner

este de asemenea un vector aleator repartizat multinomial cu parametrii $n = n_1 + \ldots + n_s$ și $\mathbf{p}$. Într-adevăr, funcția caracteristică a lui $\tilde{\mathbf{x}}$ este

$$\varphi_{\tilde{\mathbf{x}}}(\mathbf{t}) = \prod_{j=1}^{s} \varphi_{Mu}(\mathbf{t} \mid n_j, \mathbf{p}) = \prod_{j=1}^{s} (p_0 + p_1 e^{it_1} + \ldots + p_k e^{it_k})^{n_j} =$$

$$= (p_0 + p_1 e^{it_1} + \ldots + p_k e^{it_k})^n = \varphi_{Mu}(\mathbf{t} \mid n, \mathbf{p}),$$

de unde se deduce imediat rezultatul căutat.

Funcția caracteristică se poate utiliza de asemenea pentru determinarea momentelor de diverse ordine ale unui vector aleator k—dimensional $\tilde{\mathbf{x}} \in Mu(n, \mathbf{p})$. Se obțin rezultatele următoare:

$$\begin{aligned} M(\tilde{x}_j) &= np_j; \\ M(\tilde{x}_i \tilde{x}_j) &= n(n-1)p_i p_j, \quad i \neq j; \\ M(\tilde{x}_j^2) &= n^2 p_j^2 + np_j(1-p_j), \end{aligned} \tag{12.4.12}$$

unde $\tilde{x}_j$, $1 \leqslant j \leqslant k$ sînt componentele vectorului $\tilde{\mathbf{x}}$.

Momentele centrate se obțin ușor determinînd momentele variabilelor aleatoare

$$\tilde{y}_j = \tilde{x}_j - M(\tilde{x}_j) = \tilde{x}_j - m_j.$$

Obținem astfel

$$\begin{aligned} M(\tilde{y}_i \tilde{y}_j) &= -np_i p_j, \quad i \neq j, \\ M(\tilde{y}_j^2) &= np_j(1-p_j), \end{aligned} \tag{12.4.13}$$

de unde deducem

$$\begin{aligned} V(\tilde{x}_j) &= np_j(1-p_j), \\ \operatorname{cov}(\tilde{x}_j, \tilde{x}_i) &= -np_j p_i, \quad i \neq j. \end{aligned} \tag{12.4.14}$$

Printr-un raționament analog celui făcut în cazul repartiției binomiale se poate arăta că vectorul aleator

$$\left(\frac{\tilde{x}_1 - np_1}{np_1(1-p_1)}, \ldots, \frac{\tilde{x}_k - np_k}{np_k(1-p_k)} \right)'$$

converge în repartiție spre un vector normal standard.

Scanned with OKEN Scanner

Să considerăm un vector aleator $\tilde{\mathbf{x}} \in Mu(n, \mathbf{p})$ și fie $\tilde{\mathbf{x}} = (\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2)$ o partiție a sa, unde $\tilde{\mathbf{x}}_1$ are k_1 componente cu $0 < k_1 < k$, și anume $\tilde{\mathbf{x}}_1 = (\tilde{x}_{j_1}, \ldots, \tilde{x}_{j_{k_1}})'$. Dacă $(\mathbf{p}_1, \mathbf{p}_2)$ este partiția corespunzătoare a lui $\mathbf{p}$, atunci se poate arăta fără dificultate că $\tilde{\mathbf{x}}_1 \in Mu(n, \mathbf{p}_1)$ și $\tilde{\mathbf{x}}_2 \in Mu(n, \mathbf{p}_2)$. Mai mult, funcția de masă a lui $\tilde{\mathbf{x}}_2$ condiționată de $\tilde{\mathbf{x}}_1 = \mathbf{r}_1$ este $f_{Mu}(\mathbf{x}_2 \mid n_{2|1}(\mathbf{r}_1), \mathbf{p}_{2|1})$, unde

$$n_{2|1}(\mathbf{r}_1) = n - (r_{j_1} + \ldots + r_{j_{k_1}}) \tag{12.4.15}$$

și

$$\mathbf{p}_{2|1} = \frac{1}{(1 - (p_{j_1} + \ldots + p_{j_{k_1}}))} \mathbf{p}_2. \tag{12.4.16}$$

Exemplu. Să considerăm un vector aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2)$ care are o repartiție multinomială corespunzătoare cazului $k = 2$ (numită și *repartiție trinomială*) de parametrii $n = 3$ și $\mathbf{p} = (1/5, 3/5)'$. Avem atunci $p_0 = 1 - (p_1 + p_2) = 1/5$ și deci funcția de masă corespunzătoare este

$$f_{Mu}(r_1, r_2 \mid 3, (1/5, 3/5)) = \frac{3!}{r_0!\, r_1!\, r_2!} p_0^{r_0} p_1^{r_1} p_2^{r_2},$$

unde $r_0, r_1, r_2 \in \{0, 1, 2, 3\}$ cu proprietatea că $r_0 + r_1 + r_2 = 3$ adică

$$f_{Mu}(r_1, r_2 \mid 3, (1/5, 3/5)) = \frac{3!}{(3 - r_1 - r_2)!\, r_1!\, r_2!} \left(\frac{1}{5}\right)^{1 - r_1 - r_2} \cdot \left(\frac{1}{5}\right)^{r_1} \cdot \left(\frac{3}{5}\right)^{r_2}$$

pentru toți $r_1, r_2 \in \{0, 1, 2, 3\}$ pentru care $r_1 + r_2 \leqslant 3$. Valorile acestei funcții de masă sînt date în tabelul 12.4.1

Tabelul 12.4.1

r_1	r_2				Suma pe linii
	0	1	2	3	
0	1/125	9/125	27/125	27/125	64/125
1	3/125	18/125	27/125	0	48/125
2	3/125	9/125	0	0	12/125
3	1/125	0	0	0	1/125
Suma pe coloane	8/125	36/125	54/125	27/125	1

Momentele și momentele centrate ale lui $\tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2)'$ sînt:

$$M(\tilde{x}_1) = np_1 = 3/5; \quad M(\tilde{x}_2) = np_2 = 9/5;$$

$$M(\tilde{x}_1\tilde{x}_2) = n(n-1)p_1p_2 = 18/25;$$

$$M(\tilde{x}_1^2) = n^2p_1^2 = 21/25; \quad M(\tilde{x}_2^2) = n^2p_2^2 = 99/25$$

și

$$V(\tilde{x}_1) = np_1(1-p_1) = 12/25; \quad V(\tilde{x}_2) = np_2(1-p_2) = 18/25;$$

$$\text{cov}\,(\tilde{x}_1, \tilde{x}_2) = -np_1p_2 = -9/25.$$

Matricea de covarianță a lui $\tilde{\mathbf{x}}$ este deci

$$V(\tilde{\mathbf{x}}) = \begin{pmatrix} 12/25 & -9/25 \\ -9/25 & 18/25 \end{pmatrix}.$$

Se poate vedea totodată că funcția de masă a variabilei aleatoare $\tilde{x}_1$ este (vezi ultima coloană din tabelul 12.4.1):

$$f_{\tilde{x}_1}(r_1) = \begin{cases} 64/125 & \text{pentru } r_1 = 0, \\ 48/125 & \text{pentru } r_1 = 1, \\ 12/125 & \text{pentru } r_1 = 2, \\ 1/125 & \text{pentru } r_1 = 3 \end{cases}$$

și coincide cu $f_b(r_1 \mid 3, 1/5)$, cum era de așteptat, conform observațiilor făcute mai sus.

Analog se poate vedea că $f_{\tilde{x}_2}(r_2) = f_b(r_2 \mid 3, 3/5)$.

În sfîrșit, funcția de masă a lui $\tilde{x}_2$ condiționată de $\tilde{x}_1 = r_1$ este $f_b(\cdot \mid n_{2|1}(r_1), p_{2|1})$, unde

$$n_{2|1}(r_1) = 3 - r_1, \quad p_{2|1} = p_2/(1-p_1) = 3/4.$$

Prin urmare obținem

$$f(r_2 \mid r_1) = f_b(r_2 \mid 3-r_1, 3/4).$$

De exemplu, pentru $r_1 = 0$ obținem valorile:

$$f_b(0 \mid 3, 3/4) = 1/64; \quad f_b(1 \mid 3, 3/4) = 9/64;$$

$$f_b(2 \mid 3, 3/4) = 27/64; \quad f_b(3 \mid 3, 3/4) = 27/64,$$

care se puteau dealtfel obține direct din tabelul 12.4.1, împărțind numerele din prima linie a tabelului la suma lor.

Scanned with OKEN Scanner

12.5. PROCESUL HIPERGEOMETRIC

În esență, procesul hipergeometric este un proces de generare de date analog cu procesul Bernoulli, dar în care variabilele aleatoare nu mai sînt nici independente nici identic repartizate.

Pentru o prezentare mai simplă a procesului, vom considera cazul particular al unei urne conținînd N bile în total, dintre care R roșii și $N-R$ albe. Se fac extrageri succesive din urnă, fără a repune bila extrasă în urnă. Dacă $\tilde{x}_n$ este variabila aleatoare care ia valoarea 1 sau 0 după cum apare o bilă roșie sau albă la extragerea de rang n, atunci este clar că variabilele aleatoare $\tilde{x}_1, \tilde{x}_2, \ldots$ astfel introduse nu mai sînt independente.

Funcția de masă a lui $\tilde{x}_1$ este evident

$$f_{\tilde{x}_1}(x_1)=\begin{cases}\dfrac{R}{N} & \text{dacă } x_1=1,\\[2ex] \dfrac{N-R}{N} & \text{dacă } x_1=0,\\[2ex] 0 & \text{în celelalte cazuri,}\end{cases}$$

sau încă

$$f_{\tilde{x}}(x_1)=\begin{cases}\dfrac{A_R^{x_1}\cdot A_{N-R}^{1-x_1}}{A_N^1} & \text{dacă } x_1\in\{0,1\},\\[2ex] 0 & \text{dacă } x_1\notin\{0,1\},\end{cases} \qquad (12.5.1)$$

unde

$$A_n^r=\frac{n!}{(n-r)!}$$

reprezintă *aranjamente de n luate cîte r* și este definit pentru orice n și orice r, $0\leqslant r\leqslant n$.

Pentru a calcula funcția de masă a lui $\tilde{x}_{n+1}$ condiționată de $\tilde{x}_1=x_1, \ldots, \tilde{x}_n=x_n$, să observăm mai întîi că

$$r=x_1+\ldots+x_n$$

reprezintă numărul bilelor roșii extrase în cele n extrageri și deci $n-r$ reprezintă numărul bilelor albe extrase. Este clar că

Scanned with OKEN Scanner

trebuie să avem $0 \leqslant r \leqslant R$ și $0 \leqslant n - r \leqslant N - R$, de unde se deduce imediat condiția

$$\max(0, n - N + R) \leqslant r \leqslant \min(R, n). \tag{12.5.2}$$

Cu alte cuvinte, vor trebui considerate ca posibile numai valorile $(x_1, \ldots, x_n)$ pentru care sînt satisfăcute condițiile

$$\begin{cases} x_j \in \{0, 1\}, \quad 1 \leqslant j \leqslant n, \\ \max(0, n - N + R) \leqslant \sum_{j=1}^{n} x_j \leqslant \min(R, n). \end{cases} \tag{12.5.3}$$

Să observăm de asemenea că numărul total al bilelor rămase în urnă după extragerea de rang n este $N-n$, dintre care $R-r$ sînt roșii și $N-R-(n-r)$ albe.

Rezultă atunci că pentru orice $(x_1, \ldots, x_n)$ care verifică condițiile (12.5.3), funcția de masă a lui $\tilde{x}_{n+1}$ condiționată de $\tilde{x}_1 = x_1, \ldots, \tilde{x}_n = x_n$ este

$$f(x_{n+1} \mid x_1, \ldots, x_n) = \frac{A_{R-r}^{x_{n+1}} \cdot A_{N-R-n+r}^{1-x_{n+1}}}{A_{N-n}^{1}} \tag{12.5.4}$$

dacă $x_{n+1} \in \{0, 1\}$ și nulă în caz contrar.

Revenind la cazul general, vom spune că $\tilde{x}_1, \tilde{x}_2, \ldots$ este un *proces generator de date hipergeometric cu parametrii* R *și* N unde $0 < R < N$ sînt întregi, dacă funcția de masă a lui $\tilde{x}_1$ este dată de (12.5.1), iar funcția de masă a lui $\tilde{x}_{n+1}$ condiționată de $\tilde{x}_1 = x_1, \ldots, \tilde{x}_n = x_n$ este dată de (12.5.4) pentru orice $(x_1, \ldots, x_n)$ care verifică condițiile (12.5.3).

Pentru a obține funcția de masă a vectorului aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n)$, să observăm că

$$f_{\tilde{\mathbf{x}}}(x_1, \ldots, x_n) = f_{\tilde{x}_1}(x_1) \prod_{i=2}^{n} f_{\tilde{x}_i}(x_i \mid x_1, \ldots, x_{i-1})$$

Utilizînd (12.5.1) și (12.5.4) din formula de mai sus obținem

$$f_{\tilde{\mathbf{x}}}(x_1, \ldots, x_n) = \begin{cases} \dfrac{A_R^r \cdot A_{N-R}^{n-r}}{A_N^n} & \text{pentru orice } r \text{ întreg care} \\ & \text{verifică condiția (12.5.2),} \\ 0 & \text{în caz contrar.} \end{cases} \tag{12.5.5}$$

Din (12.5.5) reiese că $(\tilde{r}, \tilde{n})$ este o statistică suficientă pentru (R, N).

12.5.1. REPARTIȚIA HIPERGEOMETRICĂ

Ca și în cazul procesului Bernoulli, vom considera mai întîi un proces de oprire care constă în fixarea în prealabil a numărului de observații n.

Este clar că

$$\tilde{x}_{(n)} = \tilde{x}_1 + \ldots + \tilde{x}_n \tag{12.5.6}$$

reprezintă numărul de bile roșii care vor apare în cele n extrageri (numărul de valori egale cu 1, în general). Rezultă că

$$P(\tilde{x}_{(n)} = r) = \Sigma f_{\tilde{x}}(x_1, \ldots, x_n),$$

unde suma se face în raport cu toate valorile $(x_1, \ldots, x_n)$ pentru care $x_1 + \ldots + x_n = r$. Cum $f_x(x_1, \ldots, x_n)$ este constantă pe mulțimea acestor valori și există C_n^r valori $(x_1, \ldots, x_n)$ de acest tip, din (12.5.5) și (12.5.6) rezultă

$$P(\tilde{x}_{(n)} = r) = C_n^r \left(\frac{A_R^r \cdot A_{N-R}^{n-r}}{A_N^n} \right)$$

pentru orice întreg r care verifică condițiile (12.5.2) și nulă în celelalte cazuri. Deoarece

$$\frac{C_n^r \cdot A_R^r \cdot A_{N-R}^{n-r}}{A_N^n} = \frac{n!}{r!(n-r)!} \cdot \frac{R!}{(R-r)!} \cdot \frac{(N-R)!}{(N-R-n+r)!} \cdot \frac{(N-n)!}{N!} =$$

$$= \frac{R!}{r!(R-r)!} \cdot \frac{(N-R)!}{(n-r)!(N-R-n+r)!} \cdot \frac{n!(N-n)!}{N!}$$

rezultă că putem scrie

$$P(\tilde{x}_{(n)} = r) = \frac{C_R^r \cdot C_{N-R}^{n-r}}{C_N^n} \tag{12.5.7}$$

pentru orice întreg r care verifică (12.5.2) și nulă în celelalte cazuri.

Se pune în evidență în acest mod o repartiție pe care o vom numi *repartiția hipergeometrică*. Mai precis, vom spune că o variabilă aleatoare urmează o *repartiție hipergeometrică de*

Scanned with OKEN Scanner

parametrii n, R și N, unde $N \geqslant \max(n, R)$, dacă funcția sa de masă, pe care o notăm prin $f_h(r \mid n, R, N)$ este de forma

$$f_h(r \mid n, R, N) = \begin{cases} \dfrac{C_R^r \mid C_{N-R}^{n-r}}{C_N^n} & \text{pentru orice întreg } r \\ & \text{care verifică (12.5.2),} \\ 0 & \text{în celelalte cazuri.} \end{cases} \tag{12.5.8}$$

Din cele de mai sus rezultă că numărul de apariții ale bilei roșii în n extrageri fără repunere dintr-o urnă care conține R bile roșii și $N-R$ bile albe este o variabilă aleatoare care are o repartiție hipergeometrică de parametrii n, R și N. Mulțimea variabilelor aleatoare cu această repartiție va fi notată prin $H(n, R, N)$.

Pentru a determina momentele de diverse ordine vom face calcule directe, funcția caracteristică fiind în acest caz greu de obținut (variabilele aleatoare $\tilde{x}_1, \ldots, \tilde{x}_n$ nu mai sînt independente!). Dacă $\tilde{x} \in H(n, R, N)$, atunci

$$M(\tilde{x}) = \sum_{r=0}^{n} r f_h(r \mid n, R, N) = \sum_{r=1}^{n} r f_h(r \mid n, R, N) = \tag{12.5.9}$$

$$= (nR/N) \sum_{r=0}^{n-1} f_h(r \mid n-1, R-1, N-1) = nR/N.$$

Analog, pentru calculul dispersiei, avem mai întîi

$$M(\tilde{x}^2) - M(\tilde{x}) = \sum_{r=0}^{n} (r^2 - r) f_h(r \mid n, R, N) =$$

$$= \sum_{r=2}^{n} r(r-1) f_h(r \mid n, R, N) =$$

$$= \frac{n(n-1) R(R-1)}{N(N-1)} \sum_{r=0}^{n-2} f_h(r \mid n-2, R-2, N-2) = \frac{n(n-1) R(R-1)}{N(N-1)}$$

și deci

$$M(\tilde{x}^2) = \frac{n(n-1) R(R-1)}{N(N-1)} + \frac{nR}{N}. \tag{12.5.10}$$

Rezultă atunci că dispersia este

$$V(\tilde{x}) = \left(\frac{N-n}{N-1}\right) n \frac{R}{N}\left(1 - \frac{R}{N}\right). \qquad (12.5.11)$$

Deoarece $p = R/N$ reprezintă proporția de bile roșii din urnă, rezultă că raportul dintre dispersia unei variabile aleatoare din clasa $H(n, R, N)$ și dispersia unei variabile aleatoare din clasa $B(n, p = R/N)$ este $(N-n)/(N-1)$, care este subunitar pentru orice $n \in \{2, ..., N\}$. Dispersia în extragerile fără înlocuire este deci mai mică decît dispersia în extragerile cu înlocuire din aceeași urnă.

12.5.2. REPARTIȚIA PASCAL-HIPERGEOMETRICĂ

Să considerăm pentru procesul hipergeometric cu parametrii R și N un proces de oprire caracterizat de cerința ca numărul bilelor roșii extrase să fie un număr fixat în prealabil r. Rezultă atunci că numărul de extrageri necesar pentru extragerea a r bile roșii este o variabilă aleatoare $\tilde{n}$.

Dacă experimentul se termină la extragerea n, atunci $x_n = 1$, iar numărul valorilor posibile $(x_1, ..., x_{n-1}, 1)$ care au exact r valori egale cu 1 (bile roșii) este C_{n-1}^{r-1}, așa cum am văzut și în cazul procesului Bernoulli. Funcția de masă a lui $\tilde{n}$ este notată prin $f_{Pah}(n \mid r, R, N)$ și este dată de relația

$$f_{Pah}(n \mid r, R, N) = \begin{cases} \dfrac{C_{n-1}^{r-1} \cdot C_{N-n}^{R-r}}{C_N^R} & \text{pentru } n \in \{r, r+1, ..., N-R+r\}, \\ 0 & \text{în rest,} \end{cases} \qquad (12.5.12)$$

unde $r \leqslant R$. Dacă $r > R$, atunci toate cele N bile sînt extrase și deci probabilitatea ca numărul extragerilor să fie N este egală cu 1.

Scanned with OKEN Scanner

Dacă o variabilă aleatoare $\tilde{x}$ are funcția de masă dată de (12.5.12), spunem ca are o *repartiție Pascal-hipergeometrică* de parametrii r, R și N. Mulțimea variabilelor aleatoare care au repartiție Pascal hipergeometrică de parametrii r, R, N va fi notată prin $Pah\,(r, R, N)$.

Din cele de mai sus rezultă că numărul de bile extrase fără repunere dintr-o urnă care conține R bile roșii și $N-R$ bile albe cu scopul de a obține r bile roșii este o variabilă aleatoare cu repartiție Pascal-hipergeometrică de parametrii r, R și N.

Pentru a calcula momentele unei variabile aleatoare $\tilde{x} \in Pah(r, R, N)$ vom face din nou calcule directe. Avem mai întîi

$$M(\tilde{x}) = \sum_{n=r}^{N-R+r} n f_{Pah}(n \mid r, R, N) =$$

$$= \frac{r(N+1)}{(R+1)} \sum_{n=r}^{N-R+r} \frac{C_n^r \cdot C_{N-n}^{R-r}}{C_{N+1}^{R+1}}.$$

Dacă notăm $N+1 = N'$, $R+1 = R'$, $n+1 = n'$ și $r+1 = r'$, relația precedentă devine

$$M(\tilde{x}) = \frac{r(N+1)}{(R+1)} \sum_{n'=r'}^{N'-R'+r'} f_{Pah}(n' \mid r', R', N')$$

și deci

$$M(\tilde{x}) = \frac{r(N+1)}{(R+1)}. \qquad (12.5.13)$$

Procedînd analog obținem

$$V(\tilde{x}) = \frac{r(N+1)(N-R)(R-r+1)}{(R+1)^2(R+2)}. \qquad (12.5.14)$$

Comparînd (12.5.13) cu (12.5.19) se observă că este de așteptat ca numărul de bile care trebuie extrase pentru a obține r bile roșii dintr-o urnă care conține bile roșii și albe este mai mare cînd extragerile se fac cu repunere față de cazul cînd extragerile se fac fără repunere.

Scanned with OKEN Scanner

12.5.3. REPARTIȚIA HIPERGEOMETRICĂ NEGATIVĂ

Să considerăm un proces hipergeometric de parametrii R și N și același proces de oprire considerat în § 12.5.2. Numărul de bile albe care se extrag înainte de a obține r bile roșii este o variabilă aleatoare $\tilde{z}$ a cărei funcție de masă este notată prin $f_{hn}(z \mid r, R, N)$ și este dată de relația

$$f_{hn}(z \mid r, R, N) = \begin{cases} \dfrac{C_{z+r-1}^{r-1} \cdot C_{N-z-r}^{R-r}}{C_N^R} & \text{pentru } z \in \{0, 1, \ldots, N-R\} \\ 0 & \text{în rest,} \end{cases} \tag{12.5.15}$$

unde $r \leqslant R$.

Vom spune că o variabilă aleatoare $\tilde{x}$ are *repartiție hipergeometrică negativă de parametrii* r, R și N dacă funcția sa de masă este dată de relația (12.5.15). Mulțimea variabilelor aleatoare cu această repartiție este notată prin $Hn(r, R, N)$.

Din cele de mai sus rezultă că numărul de bile albe care apar în extrageri fără repunere dintr-o urnă care conține R bile roșii și $N-R$ bile albe înainte de a obține r bile roșii este o variabilă aleatoare din clasa $Hn(r, R, N)$.

Este clar că dacă $\tilde{x} \in Hn(r, R, N)$, atunci $\tilde{x}+r \in Pah\ (r, R, N)$ și deci, ținînd seama de (12.5.13) și (12.5.14) obținem

$$M(\tilde{x}) = M(\tilde{x}+r) - r = \frac{r(N+1)}{R+1} - r,$$

adică

$$M(\tilde{x}) = \frac{r(N-R)}{R+1}. \tag{12.5.16}$$

Analog, avem

$$V(\tilde{x}) = V(\tilde{x}+r) = \frac{r(N+1)(N-R)(R-r+1)}{(R+1)^2(R+2)}. \tag{12.5.17}$$

Scanned with OKEN Scanner

12.5.4. EXEMPLE

Dacă n este mare în raport cu R și $N-R$, atunci diferența dintre $f_h(r \mid n, R, N)$ și $f_b(r \mid n, p)$, unde $p = R/N$, este foarte pronunțată, așa cum rezultă și din exemplul următor.

O urnă conține $N = 3$ bile, dintre care $R = 1$ bilă roșie și $N-R = 2$ bile albe. Se fac $n = 2$ extrageri fără repunere. Funcția de masă corespunzătoare $f_h(r \mid 2, 1, 3)$ are valori nenule numai pentru $r = 0$ și $r = 1$:

$$f_h(0 \mid 2, 1, 3) = 1/3; \quad f_h(1 \mid 2, 1, 3) = 2/3.$$

Dacă din această urnă se fac două extrageri *cu înlocuire*, cum $p = R/N = 1/3$, rezultă că $f_b(r \mid 2, 1/3)$ are valorile nenule următoare:

$$f_b(0 \mid 2, 1/3) = 4/9; \quad f_b(1 \mid 2, 1/3) = 4/9; \quad f_b(2 \mid 2, 1/3) = 1/9$$

și diferența dintre valorile celor două funcții de masă este evidentă.

Dacă însă n este mic în raport cu R și $N-R$ și $p = R/N$, atunci diferența dintre $f_b(r \mid n, p)$ și $f_h(r \mid n, R, N)$ este foarte mică, după cum se poate observa și din tabelul 12.5.1, care conține valorile acestor funcții de masă pentru $R = 250$, $N = 1000$ și $n = 4$ (evident, $p = 1/4$).

Tabelul 12.5.1

r	$f_h(r \mid 4, 250, 1000)$	$f_b(r \mid 4, 1/4)$
0	0,3158	0,3164
1	0,4227	0,4219
2	0,2111	0,2109
3	0,0466	0,0469
4	0,0038	0,0039

Scanned with OKEN Scanner

12.6. PROCESUL HIPERGEOMETRIC MULTIDIMENSIONAL

După cum am văzut, procesul hipergeometric este analogul procesului Bernoulli în cazul cînd se înlocuiește procedeul de extragere cu repunere prin procedeul de extragere fără repunere dintr-o urnă care conține bile de două culori.

Analogul procesului Bernoulli multidimensional este procesul hipergeometric multidimensional. Pentru simplitatea limbajului, vom folosi o urnă care conține bile de $k+1$ culori, notate prin $c_0, c_1, \ldots, c_k$, numărul bilelor de culoare c_i fiind notat prin R_i, $0 \leqslant i \leqslant k$; numărul total al bilelor din urnă este evident

$$N = R_0 + R_1 + \ldots + R_k. \tag{12.6.1}$$

Rezultatul unei extrageri din această urnă va fi notat prin vectorii k—dimensionali $\mathbf{e}_0 = (0, \ldots, 0)'$, $\mathbf{e}_1 = (1; 0, \ldots, 0)', \ldots$ $\ldots, \mathbf{e}_k = (0, \ldots, 0, 1)'$, după cum apare o bilă de culoare $c_0, c_1, \ldots, c_k$ respectiv.

Efectuăm extrageri fără repunere din această urnă și notăm prin $\tilde{\mathbf{x}}_n$ vectorul aleator definit pentru extragerea de rang n prin

$$\tilde{\mathbf{x}}_n = \begin{cases} \mathbf{e}_0 & \text{dacă apare o bilă de culoare } c_0, \\ \mathbf{e}_1 & \text{dacă apare o bilă de culoare } c_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \mathbf{e}_k & \text{dacă apare o bilă de culoare } c_k. \end{cases} \tag{12.6.2}$$

Este clar că $\tilde{\mathbf{x}}_n$ reprezintă rezultatul extragerii de rang n, vectorii $\mathbf{e}_0, \mathbf{e}_1, \ldots, \mathbf{e}_k$ reprezentînd valorile posibile ale acestor vectori aleatori. Evident, vectorii aleatori $\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_n, \ldots$ astfel definiți nu sînt independenți.

Vom spune că vectorii aleatori $\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \ldots$ definiți mai sus formează un *proces hipergeometric k—dimensional de parametrii* $\mathbf{R} = (R_1, \ldots, R_k)'$ și N, unde $R_0 = N-(R_1 + \ldots + R_k)$, conform cu (12.6.1).

Se observă ușor că funcția de masă a vectorului aleator $\tilde{\mathbf{x}}_1$ (rezultatul primei extrageri) este de forma

$$f_{\tilde{\mathbf{x}}_1}(\mathbf{x}) = \begin{cases} \dfrac{A_{R_0}^{x_0} \cdot A_{R_1}^{x_1} \dots A_{R_k}^{x_k}}{A_N^1} & \text{dacă } \mathbf{x} = (x_1, \dots, x_k) \text{ este} \\ & \text{o valoare posibilă,} \\ 0 & \text{în caz contrar,} \end{cases} \tag{12.6.3}$$

unde

$$x_0 = 1 - (x_1 + \dots + x_k) \tag{12.6.4}$$

și

$$R_0 = N - (R_1 + \dots + R_k). \tag{12.6.5}$$

Rezultatul obținut la extragerea de rang i îl vom înregistra sub forma unui vector $\mathbf{x}_i = (x_{1i}, \dots, x_{ki})'$ ale cărui valori pot fi valorile posibile $\mathbf{e}_0, \dots, \mathbf{e}_k$. Dacă notăm

$$x_{0i} = 1 - (x_{1i} + \dots + x_{ki}), \tag{12.6.6}$$

atunci este clar că $x_{ji} = 1$ dacă și numai dacă la extragerea de rang i am obținut o bilă de culoarea c_j, $0 \leqslant j \leqslant k$.

Evident, după efectuarea a n extrageri ale căror rezultate au fost vectorii $\mathbf{x}_1, \dots, \mathbf{x}_n$, urna mai conține $N-n$ bile, dintre care

$$R_j(n) = R_j - \sum_{i=1}^{n} x_{ji}, \quad 0 \leqslant j \leqslant k, \tag{12.6.7}$$

sînt de culoarea c_j, $0 \leqslant j \leqslant k$.

Funcția de masă a vectorului aleator $\tilde{\mathbf{x}}_{n+1}$ condiționată de $\tilde{\mathbf{x}}_1 = \mathbf{x}_1, \dots, \tilde{\mathbf{x}}_n = \mathbf{x}_n$ este atunci dată de relația

$$f(\mathbf{x}_{n+1} \mid \mathbf{x}_1, \dots, \mathbf{x}_n) = \frac{\prod_{j=0}^{k} A^{x_{n+1,j}}}{A_{N-n}^1} \tag{12.6.8}$$

pentru orice valoare posibilă a lui $\mathbf{x}_{n+1}$.

Este clar că relațiile (12.6.3) și (12.6.8) caracterizează un proces hipergeometric k-dimensional de parametrii $\mathbf{R}$ și N, putînd fi adoptate în cazul general drept definiție a acestui proces.

Scanned with OKEN Scanner

Putem acum obține funcția de masă a vectorului aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, ..., \tilde{x}_n)'$ pe calea obișnuită:

$$f(x_1, ..., x_n) = f(x_1) \prod_{i=2}^{n} f(x_i \mid x_1, ..., x_{i-1}).$$

După calcule simple se obține că valorile nenule ale funcției de masă sînt de forma

$$f(x_1, ..., x_n) = \frac{\prod_{j=0}^{k} A_{R_j}^{r_j}}{A_N^n}, \tag{12.6.9}$$

unde r_j, $0 \leqslant j \leqslant k$, satisfac condițiile

$$\begin{cases} r_j \in \{0, 1, ..., \min(N, R_j)\}, \\ \sum_{j=1}^{k} r_j \geqslant n - \left(N - \sum_{j=1}^{k} R_j\right) \end{cases} \tag{12.6.10}$$

și sînt definiți de relațiile

$$r_j = R_j - R_j(n) = \sum_{i=1}^{n} x_{ji}, \quad 0 \leqslant j \leqslant k, \tag{12.6.11}$$

iar N, $R_0, ..., R_k$ sînt întregi pozitivi care verifică condițiile:

$$\max(n, R_0, R_1, ..., R_k) \leqslant N \tag{12.6.12}$$

și

$$R_0 + R_1 + ... + R_k = N. \tag{12.6.13}$$

Din aceste condiții rezultă că, notînd prin $\mathbf{r}$ vectorul cu componentele $r_1, ..., r_k$, statistica $(\tilde{\mathbf{r}}, \tilde{n})$ este suficientă pentru $(\mathbf{R}, N)$.

Scanned with OKEN Scanner

12.6.1. REPARTIȚIA HIPERGEOMETRICĂ MULTIDIMENSIONALĂ

Ca și în cazul unidimensional, vom considera un proces de oprire care constă în fixarea în prealabil a numărului n de observații și vom căuta să determinăm repartiția vectorului aleator

$$\tilde{\mathbf{x}}_{(n)} = \tilde{\mathbf{x}}_1 + \ldots + \tilde{\mathbf{x}}_n. \tag{12.6.14}$$

Pentru orice $\mathbf{r} \in R^k$ avem

$$f_{\mathbf{x}_{(n}}(\mathbf{r}) = P(\tilde{\mathbf{x}}_1 + \ldots + \tilde{\mathbf{x}}_n = \mathbf{r}) = \sum f(\mathbf{x}_1, \ldots, \mathbf{x}_n),$$

unde suma se face pentru toate valorile $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ pentru care $\mathbf{x}_1 + \ldots + \mathbf{x}_n = \mathbf{r}$. După calcule simple se obține

$$f_{\tilde{\mathbf{x}}_{(n)}}(\mathbf{r}) = \frac{C_{R_0}^{r_0} \cdot C_{R_1}^{r_1} \ldots C_{R_k}^{r_k}}{C_N^n}, \tag{12.6.15}$$

unde N, $R_0, \ldots, R_k$ sînt întregi pozitivi care verifică condițiile (12.6.12) și (12.6.13), iar r_j, $0 \leqslant j \leqslant k$, satisfac condițiile (12.6.10).

Vom spune că un vector aleator k-dimensional are o *repartiție hipergeometrică de parametrii* n, $\mathbf{R}$ *și* N, dacă funcția sa de masă, pe care o vom nota prin $f_h(\mathbf{r} \mid n, \mathbf{R}, N)$ este dată de relația

$$f_h(\mathbf{r} \mid n, \mathbf{R}, N) = \begin{cases} \dfrac{C_{R_0}^{r_0} \cdot C_{R_1}^{r_1} \ldots C_{R_k}^{r_k}}{C_N^n} & \text{dacă } \mathbf{r} \text{ verifică condițiile (12.6.10),} \\ 0 & \text{în caz contrar,} \end{cases} \tag{12.6.16}$$

unde parametrii n, $\mathbf{R}$ și N verifică condițiile (12.6.11), (12.6.12) și (12.6.13). Mulțimea vectorilor aleatori avînd această repartiție va fi notată prin $H(n, \mathbf{R}, N)$.

Să considerăm un vector aleator $\tilde{\mathbf{x}} \in H(n, \mathbf{R}, N)$ și să partiționăm acest vector în vectorii $\tilde{\mathbf{x}}_1 = (\tilde{x}_{i_1}, \ldots, \tilde{x}_{i_{k_1}})'$ cu $0 < k_1 < k$ și $\tilde{\mathbf{x}}_2$. Dacă se efectuează partiții analoage ale vectorilor $\mathbf{R}$ și $\mathbf{r}$ în vectorii $\mathbf{R}_1$ și $\mathbf{R}_2$ și respectiv $\mathbf{r}_1$ și $\mathbf{r}_2$, atunci rezultă destul de simplu că $\tilde{\mathbf{x}}_1 \in H(n, \mathbf{R}_1, N)$ și $\tilde{\mathbf{x}}_2 \in H(n, \mathbf{R}_2, N)$. Mai mult, se arată că funcția de masă a lui $\tilde{\mathbf{x}}_2$ condiționată de $\tilde{\mathbf{x}}_1 = \mathbf{r}_1$ este

$$f_h(\mathbf{x}_2 \mid n_{2|1}(\mathbf{r}_1), \mathbf{R}_2, N_{2|1}),$$

unde

$$n_{2|1}(\mathbf{r}_1) = n - \sum_{j=1}^{k_1} r_{i_j} \tag{12.6.17}$$

și

$$N_{2|1} = N - \sum_{j=1}^{k_1} R_{i_j}. \tag{12.6.18}$$

Din acest rezultat se deduce imediat că, în particular, dacă $\tilde{\mathbf{x}} \in H(n, \mathbf{R}, N)$, atunci componentele sale $\tilde{x}_j$ aparțin clasei $H(n, R_j, N)$ și deci putem obține imediat media și dispersia corespunzătoare din formulele (12.5.9) și (12.5.11):

$$M(\tilde{x}_j) = n \frac{R_j}{N}; \tag{12.6.19}$$

$$V(\tilde{x}_j) = \left(\frac{N-n}{N-1}\right) n \frac{R_j}{N}\left(1 - \frac{R_j}{N}\right). \tag{12.6.20}$$

Mai mult, prin calcul direct se poate obține

$$\operatorname{cov}(\tilde{x}_i, \tilde{x}_j) = -\frac{N-n}{N-1} \cdot \frac{nR_iR_j}{N^2}. \tag{12.6.21}$$

12.6.2. *Exemplu.* Să considerăm un vector aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2)'$ care are o repartiție hipergeometrică corespunzătoare cazului $k = 2$, de parametrii $n = 3$, $\mathbf{R} = (2,6)'$ și $N = 10$. Avem deci $R_0 = 10-2-6=2$, iar funcția de masă corespunzătoare este

$$f_h(r_1, r_2 \mid (2,6)', 10) = \frac{C_2^{r_0} \cdot C_2^{r_1} \cdot C_6^{r_2}}{C_{10}^3},$$

unde r_1 și r_2 satisfac condițiile (vezi și formulele (12.6.10)—(12.6.13)):

$$1 \leqslant r_1 + r_2 \leqslant 3;$$

$$r_1 \in \{0, 1, 2\};\ r_2 \in \{0, 1, 2, 3, 4, 5, 6\},$$

iar

$$r_0 = 3 - (r_1 + r_2).$$

Scanned with OKEN Scanner

Valorile funcției de masă $f_h(r_1, r_2 \mid 3, (2,6)', 10)$ sînt date în tabelul 12.6.1.

Tabelul 12.6.1

r_1	r_2				Suma pe linii
	0	1	2	3	
0	0	3/60	15/60	10/60	28/60
1	1/60	12/60	15/60	0	28/60
2	1/60	3/60	0	0	4/60
3	0	0	0	0	0
Suma pe coloane	2/60	18/60	30/60	10/60	1

Comparînd acest tabel cu tabelul 12.4.1, se pun în evidență diferențele dintre $f_{mu}(r_1, r_2 \mid 3, (1/5, 3/5)')$ și $f_h(r_1, r_2 \mid 3, (2,6)', 10)$, care ilustrează deosebirea dintre extragerea cu repunere și fără repunere.

Momentele și momentele centrate mai importante ale acestui vector aleator sînt:

$$M(\tilde{x}_1) = 3/5; \; M(\tilde{x}_2) = 9/5;$$

$$V(\tilde{x}_1) = 28/75; \; V(\tilde{x}_2) = 14/25,$$

și

$$\operatorname{cov}(\tilde{x}_1, \tilde{x}_2) = -7/25.$$

Urmează atunci că matricea de covarianță a vectorului aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2)'$ este

$$V(\tilde{\mathbf{x}}) = \begin{pmatrix} 28/75 & -7/25 \\ -7/25 & 14/25 \end{pmatrix}$$

Se poate de asemenea vedea că funcția de masă a variabilei aleatoare $\tilde{x}_1$ (vezi ultima coloană a tabelului 12.6.1) este:

$$f_{\tilde{x}_1}(0) = 28/60; \; f_{\tilde{x}_1}(1) = 28/60; \; f_{\tilde{x}_1}(2) = 4/60$$

și coincide cu $f_h(\cdot \mid 3, 2, 10)$, cum era de așteptat. Analog se obține că $f_{\tilde{x}_2}(r_2) = f_h(r_2 \mid 3, 6, 10)$.

Scanned with OKEN Scanner

În sfîrșit, funcția de masă a lui $\tilde{x}_2$ condiționată de $\tilde{x}_1=r_1$ este $f_h(r_2 \mid n_{2|1}(r_1), R_2, N_{2|1})$, unde

$$n_{2|1}(r_1) = n - r_1 = 3 - r_1$$

și

$$N_{2|1} = N - R_1 = 10 - 2 = 8.$$

Rezultă atunci că

$$f(r_2 \mid r_1) = f_h(r_2 \mid 3 - r_1, 6, 8).$$

De exemplu, pentru $r_1 = 0$ obținem valorile nenule

$$f_h(1 \mid 3, 6, 8)=3/28; \quad f_h(2 \mid 3, 6, 8)=15/28; \quad f_h(3|3, 6, 8)=10/28,$$

rezultat care se putea obține și direct, împărțind numerele din prima linie a tabelului 12.6.1 la suma lor.

12.7. PROCESUL POISSON

Procesul Bernoulli și procesul hipergeometric sînt procese generatoare de date care apar în selecția aleatoare din populații, cum ar fi, de exemplu, persoane, unități dintr-un lot de produse, bile dintr-o urnă etc. În cele ce urmează ne vom ocupa de un proces generator de date care descrie producerea unor anumite fenomene la momente de timp aleatoare, ca de exemplu, producerea defecțiunilor în anumite sisteme („arderea" becurilor, a tuburilor electronice etc.) sau solicitările de servicii (apelurile telefonice la o centrală, solicitările de asistență medicală la o policlinică, solicitările de aterizare la un aeroport etc.).

Momentele de timp (aleatoare) la care se produc aceste fenomene vor fi notate prin $\tilde{\tau}_0, \tilde{\tau}_1, \ldots, \tilde{\tau}_i, \ldots$, iar duratele de timp (aleatoare) dintre două produceri succesive ale fenomenelor considerate vor fi notate prin $\tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_i, \ldots$, unde $\tilde{t}_i=\tilde{\tau}_i-\tilde{\tau}_{i-1}$ pentru $i = 1, 2, \ldots$ În cele ce urmează ne vom ocupa de procesul de generare de date $\tilde{t}_1, \tilde{t}_2, \ldots$ al timpilor dintre aparițiile succesive ale fenomenelor considerate, făcînd ipoteza naturală că variabilele aleatoare $(\tilde{t}_i)_{i\geq 1}$ sînt independente și identic repartizate, avînd

densitatea de probabilitate $f(t)$ și funcția de repartiție $F(t)$. Deoarece $\tilde{\tau}_i \geqslant 0$ pentru orice i, rezultă că $F(t) = 0$ pentru orice $t \leqslant 0$.

Să considerăm că observarea procesului începe la momentul T, ultima apariție a fenomenului considerat înainte de acest moment avînd loc la momentul τ_0, adică $\tau_0 < T < \tilde{\tau}_1$. Deoarece $\tilde{t}_1 = \tilde{\tau}_1 - \tau_0$, rezultă că $\tilde{t}_1$ nu poate fi observat și deci trebuie să așteptăm pînă la momentul $\tilde{\tau}_1$ pentru a putea obține informații asupra lui $\tilde{t}_2, \tilde{t}_3, \ldots$ Timpul pe care sîntem nevoiți să-l pierdem este deci

$$\tilde{\tau} = \tilde{\tau}_1 - T = \tilde{t}_1 - (T - \tau_0) > 0,$$

deoarece $\tau_0 < T < \tilde{\tau}_1$. Probabilitatea ca acest timp să depășească o anumită valoare fixată τ se obține în modul următor

$$\begin{aligned} P(\tilde{\tau} \geqslant \tau) &= P(\tilde{\tau} > \tau \mid \tilde{\tau} > 0) = \\ &= P(\tilde{t}_1 - (T - \tau_0) \geqslant \tau \mid \tilde{t}_1 - (T - \tau_0) > 0) = \\ &= \frac{P(\tilde{t}_1 \geqslant \tau + (T - \tau_0))}{P(\tilde{t}_1 > T - \tau_0)}. \end{aligned} \tag{12.7.1}$$

Rezultă atunci că probabilitatea de a aștepta mai mult decît τ unități de timp de la momentul începerii observațiilor pînă la prima apariție a fenomenului considerat depinde în general de $T - \tau_0$, conform relației (12.7.1), situație care nu este de dorit în practică. Așa cum vom vedea mai jos, pentru procesul Poisson această situație neplăcută nu poate să apară.

Pentru introducerea procesului Poisson avem nevoie de o repartiție continuă numită *repartiția exponențială (negativă)*, pe care o introducem în continuare.

12.7.1. REPARTIȚIA EXPONENȚIALĂ NEGATIVĂ

Spunem că o variabilă aleatoare continuă $\tilde{x}$ are *repartiție exponențială (negativă)* de parametru $\lambda > 0$, dacă densitatea sa de probabilitate este

$$f_e(x \mid \lambda) = \begin{cases} 0 & \text{pentru } x \leqslant 0, \\ \lambda e^{-\lambda x} & \text{pentru } x > 0. \end{cases} \tag{12.7.2}$$

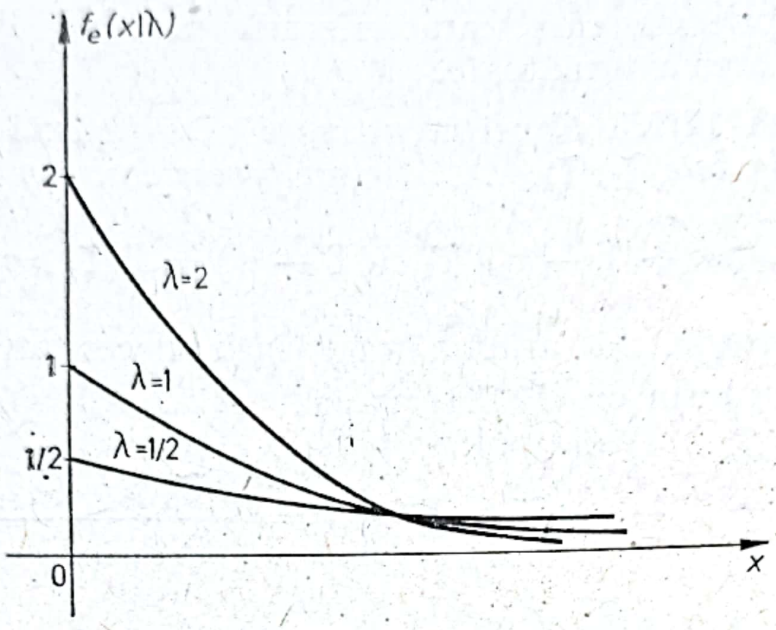


Fig. 12.7.1

Funcția de repartiție a acestei variabile aleatoare este

$$F_e(x \mid \lambda) = \begin{cases} 0 & \text{pentru } x \leqslant 0, \\ 1 - e^{-\lambda x} & \text{pentru } x > 0. \end{cases} \tag{12.7.3}$$

Mulțimea variabilelor aleatoare cu această repartiție va fi notată prin $E(\lambda)$, unde $\lambda > 0$.

Graficul funcției $f_e(x \mid \lambda)$ este prezentat în fig. 12.7.1 pentru valorile $\lambda = 1/2$, $\lambda = 1$ și $\lambda = 2$.

Dacă $\tilde{x} \in E(\lambda)$, atunci funcția sa caracteristică este

$$\varphi_e(t \mid \lambda) = \frac{\lambda}{\lambda - it}, \tag{12.7.4}$$

cu ajutorul căreia se pot calcula momentele acestei variabile aleatoare:

$$M(\tilde{x}) = 1/\lambda, \quad M_2(\tilde{x}) = 2/\lambda^2, \quad V(\tilde{x}) = 1/\lambda^2. \tag{12.7.5}$$

12.7.2. PROPRIETĂȚI ALE PROCESULUI POISSON

Vom spune că procesul generator de date $\tilde{\tau}_1, \tilde{\tau}_2, \ldots$, unde $(\tilde{\tau}_i)$ sînt variabile aleatoare nenegative, independente și identic repartizate, este un *proces Poisson de parametru* $\lambda > 0$, dacă variabilele aleatoare $(\tilde{\tau}_i)$ au repartiție exponențială de parametru $\lambda > 0$.

O proprietate care caracterizează procesul Poisson este dată în teorema următoare.

Teorema 12.7.1. *Condiția necesară și suficientă ca procesul generator de date* $\tilde{\tau}_1, \tilde{\tau}_2, \ldots$ *să fie un proces Poisson este ca*

$$P(\tilde{\tau}_1 \geqslant \tau + T - \tau_0 \mid \tilde{\tau}_1 > T - \tau_0) = P(\tilde{\tau}_1 \geqslant \tau). \quad (12.7.6)$$

Demonstrație. Să arătăm mai întîi necesitatea condiției (12.7.6). Conform cu (12.7.1) avem

$$P(\tilde{\tau}_1 \geqslant \tau + T - \tau_0 \mid \tilde{\tau}_1 > T - \tau_0) = \frac{P(\tilde{\tau}_1 \geqslant \tau + T - \tau_0)}{P(\tilde{\tau}_1 > T - \tau_0)} =$$

$$= \frac{1 - F_e(\tau + T - \tau_0)}{1 - F_e(T - \tau_0)} = \frac{e^{-\lambda(\tau + T - \tau_0)}}{e^{-\lambda(T - \tau_0)}} = e^{-\lambda\tau},$$

de unde rezultă imediat (12.7.6).

Suficiența condiției comportă calcule ceva mai lungi în demonstrație și o omitem *.

Din (12.7.6) rezultă deci că pentru un proces Poisson, probabilitatea de a aștepta mai mult de τ unități de timp de la momentul T în care încep observațiile pînă la momentul $\tilde{\tau}_1$ al primei produceri a fenomenului considerat nu depinde nici de T nici de τ_0. Cu alte cuvinte, putem începe observațiile la oricare moment T și putem considera $\tilde{\tau}_1$ ca fiind timpul scurs de la momentul T pînă la prima apariție a fenomenului considerat (vezi (12.7.1) și (12.7.6)).

Să presupunem că procedăm la observarea unui proces de parametru $\lambda > 0$ pentru o perioadă de timp fixată t și că am observat exact r produceri ale fenomenului considerat. Duratele de timp dintre producerile consecutive $t_1, t_2, \ldots, t_r$ verifică evident condiția $t_1 + t_2 + \ldots + t_r \leqslant t$. Dacă t_{r+1} este timpul scurs între producerile de rang r și $r+1$ ale fenomenului consi-

---

* Pentru demonstrație vezi Raiffa, H., Schlaifer, R. 1961.

derat, în condițiile de mai sus, densitatea de probabilitate a vectorului aleator $(\tilde{\tau}_1, \ldots, \tilde{\tau}_r, \tilde{\tau}_{r+1})'$ este

$$f(t_1, \ldots, t_r, t_{r+1} \mid t, \lambda) =$$

$$= \begin{cases} \lambda^{r+1} \exp\left(-\lambda \sum_{i=1}^{r+1} t_i\right) & \text{pentru } t_i > 0 \text{ și } \sum_{i=1}^{r} t_i \leqslant t < \sum_{i=1}^{r+1} t_i, \\ 0 & \text{în celelalte cazuri.} \end{cases} \quad (12.7.7)$$

Integrînd în raport cu t_{r+1} pe intervalul $\left(t - \sum_{i=1}^{r} t_i, +\infty\right)$, obținem ușor

$$f(t_1, \ldots, t_r \mid t, \lambda) = \begin{cases} \lambda^r e^{-\lambda t} & \text{pentru} \sum_{i=1}^{r} t_i \leqslant t, \\ 0 & \text{în rest.} \end{cases} \quad (12.7.8)$$

Prin urmare, în cazul utilizării unui proces de oprire neinformativ asupra lui λ, statistica $(\tilde{r}, \tilde{t})$ este suficientă pentru λ.

12.7.3. REPARTIȚIA POISSON

Să considerăm un proces de oprire caracterizat de faptul că durata de timp t în care se fac observațiile este fixată; evident, acesta este analogul procesului de oprire asociat procesului Bernoulli sau hipergeometric care constă în fixarea numărului n al observațiilor. Numărul $\tilde{r}$ de produceri ale fenomenului considerat pe această perioadă de timp fixată este o variabilă aleatoare ale cărei valori posibile sînt 0, 1, 2, ... Funcția de masă a acestei variabile aleatoare va fi notată prin $f(r \mid t)$. Evident, pentru $r = 0$, avem

$$f(0 \mid t) = P(\tilde{\tau}_1 > t) = e^{-\lambda t}. \quad (12.7.9)$$

Pentru $r \geqslant 1$, să considerăm intervalul $(\tau, \tau + d\tau)$ inclus în $(0, t]$ în care are loc prima producere a fenomenului. Rezultă

atunci că celelalte $r-1$ produceri au loc aproximativ în intervalul $(\tau, t]$ cu probabilitatea $f(r-1 \mid t-\tau)$ și deci putem scrie

$$f(r \mid t) = \int_0^t f(r-1 \mid t-\tau)\lambda e^{-\lambda\tau} d\tau. \qquad (12.7.10)$$

Se arată ușor prin inducție că

$$f(r \mid t) = \frac{e^{-\lambda t}(\lambda t)^r}{r!}, \qquad (12.7.11)$$

unde

$$r \in \{0, 1, 2, \ldots\}.$$

Spunem că o variabilă aleatoare discretă cu valorile posibile 0, 1, 2, ... are *repartiție Poisson de parametru* $\lambda > 0$ dacă funcția sa de masă este definită de

$$f_{Po}(r \mid \lambda) = \frac{e^{-\lambda}\lambda^r}{r!} \qquad (12.7.12)$$

pentru $r \in \{0, 1, 2, \ldots\}$ și nulă în rest.

Din cele de mai sus rezultă că numărul de produceri ale unui fenomen constatat în urma observării unui proces Poisson de parametru $\lambda > 0$, pe o perioadă de timp fixată $t > 0$, este o variabilă aleatoare cu repartiție Poisson de parametru λt.

Mulțimea variabilelor aleatoare care au repartiție Poisson de parametru $\lambda > 0$ va fi notată prin $Po(\lambda)$.

Funcția caracteristică a unei variabile aleatoare $\tilde{x} \in Po(\lambda)$ este

$$\varphi_{Po}(t \mid \lambda) = \sum_{r=0}^{\infty} e^{rti} \frac{\lambda^r e^{-\lambda}}{r!} = e^{\lambda(e^{it}-1)}. \qquad (12.7.13)$$

Se vede acum ușor că dacă $\tilde{x} \in Po(\lambda)$ și $\tilde{y} \in Po(\mu)$ sînt independente, atunci $\tilde{x}+\tilde{y} \in Po(\lambda+\mu)$. Într-adevăr, funcția caracteristică a variabilei aleatoare $\tilde{x}+\tilde{y}$ este

$$\varphi_{\tilde{x}+\tilde{y}}(t) = \varphi_{Po}(t \mid \lambda)\, \varphi_{Po}(t \mid \mu) = e^{\lambda(e^{it}-1)} \cdot e^{\mu(e^{it}-1)} =$$

$$= e^{(\lambda+\mu)(e^{it}-1)} = \varphi_{Po}(t \mid \lambda + \mu),$$

de unde rezultă imediat afirmația.

Scanned with OKEN Scanner

Utilizînd funcția caracteristică, putem determina momentele unei variabile aleatoare $\tilde{x} \in Po(\lambda)$. Obținem

$$M(\tilde{x}) = \lambda; \quad M(\tilde{x}^2) = \lambda + \lambda^2; \quad M(\tilde{x}^3) = \lambda + 3\lambda^2 + \lambda^3$$

și

$$V(\tilde{x}) = \lambda; \quad V_3(\tilde{x}) = \lambda.$$

Este interesant să observăm că media, dispersia și momentul centrat de ordinul al treilea sînt, toate, egale cu λ.

Pentru a obține repartiția limită, vom utiliza din nou funcția caracteristică a variabilei aleatoare reduse

$$\tilde{y} = \frac{\tilde{x} - M(\tilde{x})}{(V(\tilde{x}))^{1/2}} = \frac{\tilde{x} - \lambda}{\sqrt{\lambda}}$$

corespunzătoare unei variabile aleatoare $\tilde{x} \in Po(\lambda)$. Utilizînd proprietățile funcției caracteristice și rezultatul (12.7.13), obținem:

$$\varphi_{\tilde{y}}(t) = e^{-it\sqrt{\lambda}}\, \varphi_{Po}(t \mid \sqrt{\lambda}) = e^{-it\sqrt{\lambda}}\, e^{(e^{it/\sqrt{\lambda}}-1)} =$$

$$= \exp\left[\lambda(e^{it'\sqrt{\lambda}} - it/\sqrt{\lambda} - 1)\right].$$

După o dezvoltare în serie a lui $e^{it'\sqrt{\lambda}}$ obținem

$$\lim_{\lambda \to \infty} \varphi_y(t) = e^{-t^2/2},$$

care este funcția caracteristică a unei variabile normale standard.

Repartiția Poisson poate fi privită ea însăși ca o repartiție limită. Mai precis, putem arăta că

$$\lim_{n} f_b(r \mid n,\ \lambda/n) = f_{Po}(r \mid \lambda). \tag{12.7.14}$$

Într-adevăr, putem scrie

$$f_b(r \mid n, \lambda/n) = C_n^r (\lambda/n)^r (1 - \lambda/n)^{n-r} =$$

$$= \frac{\lambda^r}{r!} \cdot \frac{A_n^r}{n^r} \cdot \left(1 - \frac{\lambda}{n}\right)^{n-r}. \tag{12.7.15}$$

Deoarece

$$\lim_n \frac{A_n^r}{n^r} = \lim_n \frac{n(n-1)\dots(n-r+1)}{n^r} = 1$$

și

$$\lim_n (1 - \lambda/n)^{n-r} = e^{-\lambda},$$

relația căutată se obține imediat prin trecere la limită în (12.7.15).

12.7.4. SELECȚIA INVERSĂ ȘI REPARTIȚIA ASOCIATĂ

Să considerăm un proces Poisson de parametru $\lambda > 0$ și să-i asociem un proces de oprire caracterizat de faptul că observațiile durează pînă în momentul în care s-au produs exact r apariții ale fenomenului considerat; este vorba deci de o „selecție inversă", ca și în cazul procesului Bernoulli sau hipergeometric. Durata observațiilor este atunci o variabilă aleatoare $\tilde{\tau}$ și este definită evident de

$$\tilde{\tau} = \tilde{\tau}_1 + \dots + \tilde{\tau}_r.$$

Deoarece $\tilde{\tau}_1, \dots, \tilde{\tau}_r$ sînt variabile aleatoare independente și identic repartizate, rezultă că funcția caracteristică a lui $\tilde{\tau}$ este dată de relația

$$\varphi_{\tilde{\tau}}(t) = \prod_{j=1}^{r} \varphi_{\tilde{\tau}_j}(t) = \frac{r}{(\lambda - it)^r}. \qquad (12.7.16)$$

Vom vedea în cele ce urmează (vezi § 13.2) că funcția caracteristică dată de (12.7.16) este funcția caracteristică a unei variabile aleatoare cu *repartiție gama de parametrii* $r > 0$ și $\lambda > 0$, a cărei densitate de probabilitate este

$$f_\tau(t \mid r, \lambda) = \begin{cases} \dfrac{\lambda e^{-\lambda t}(\lambda t)^{r-1}}{\Gamma(r)} & \text{pentru } t > 0, \\ 0 & \text{pentru } t \leqslant 0, \end{cases} \qquad (12.7.17)$$

unde

$$\Gamma(r) = \int_{\infty}^{0} x^{r-1} e^{-x} dx \qquad (12.7.18)$$

este numită *funcția gama* și este definită pentru orice $r > 0$; pentru o prezentare succintă a proprietăților acestei funcții se poate vedea Anexa de la sfîrșitul lucrării (vezi § A.1).

Din cele de mai sus rezultă deci că dacă un proces Poisson de parametru $\lambda > 0$ se observă pînă în momentul cînd apar exact r produceri ale fenomenului considerat, atunci timpul total de observare $\tilde{t}$ este o variabilă aleatoare avînd repartiția gama de parametrii r și λ.

Să observăm de asemenea că avem

$$f_{\gamma}(t \mid 1, \lambda) = f_e(t \mid \lambda),$$

cum se vede ușor din (12.7.17). Repartiția exponențială este deci un caz particular al repartiției gama, care se obține pentru $r = 1$.

În sfîrșit, să observăm că repartiția gama este analoagă repartiției binomiale negative, pe care am obținut-o de asemenea considerînd o selecție inversă din procesul Bernoulli și deci repartiția exponențială este analoagă repartiției geometrice, ambele obținîndu-se drept cazuri particulare pentru $r = 1$.

Scanned with OKEN Scanner

Capitolul 13

REPARTIȚII CLASICE CONTINUE

13.1. INTRODUCERE

Vom da o prezentare succintă a principalelor repartiții clasice continue, cu excepția repartiției normale, pe care am analizat-o detaliat în capitolul 10. Accentul este pus aici în special pe prezentarea repartițiilor și nu pe procesele de generare de date în care ele ar putea să apară.

13.2. FAMILIA DE REPARTIȚII GAMA

Se spune că o variabilă aleatoare continuă are o *repartiție gama de parametrii* $r > 0$ *și* $\lambda > 0$, dacă densitatea sa de probabilitate, care este notată prin $f_\gamma(x \mid r, \lambda)$ este de forma

$$f_\gamma(x|r,\lambda) = \begin{cases} \dfrac{\lambda e^{-\lambda x}(\lambda x)^{r-1}}{\Gamma(r)} & \text{pentru } x > 0, \\ 0 & \text{pentru } x \leqslant 0. \end{cases} \tag{13.2.1}$$

Mulțimea variabilelor aleatoare care au densitatea de probabilitate $f_\gamma(x \mid r, \lambda)$ dată de (13.2.1) va fi notată prin $\gamma(r, \lambda)$, iar funcția de repartiție corespunzătoare va fi notată prin $F_\gamma(x|r, \lambda)$.

Să observăm că dacă $\tilde{x} \in \gamma(r, \lambda)$, atunci variabila aleatoare $\tilde{y} = a\tilde{x}$, unde $a \in R$, are funcția de repartiție $F_\gamma(x/a \mid r, \lambda)$, iar densitatea de probabilitate corespunzătoare este $(1/a)\, f_\gamma(x/a \mid r, \lambda)$, ceea ce arată că $a\tilde{x} \in \gamma(r, \lambda/a)$. Este atunci clar că dacă luăm $a = \lambda$, densitatea de probabilitate a lui $\lambda\tilde{x}$ nu

mai depinde de λ, fiind notată prin $f_{\gamma^*}(x \mid r)$, adică

$$f_{\gamma^*}(x \mid r) = \begin{cases} \dfrac{e^{-x} \cdot x^{r-1}}{\Gamma(r)} & \text{pentru } x > 0, \\ 0 & \text{pentru } x \leqslant 0. \end{cases} \tag{13.2.2}$$

Spunem că o variabilă aleatoare are *repartiție gama standard* dacă densitatea sa de probabilitate este dată de (13.2.2); mulțimea variabilelor aleatoare care au repartiție gama standard de parametru $r > 0$ va fi notată prin $\gamma^*(r)$. Din cele de mai sus rezultă că $\tilde{x} \in \gamma(r, \lambda)$ dacă și numai dacă $\lambda\tilde{x} \in \gamma^*(r)$. Este de asemenea util să observăm că $\gamma(r, 1) = \gamma^*(r)$ și $f_\gamma(x \mid r, 1) = f_{\gamma^*}(x \mid r)$.

Tot în legătură cu repartiția gama se introduc o serie de alte repartiții utile.

Vom spune că o variabilă aleatoare are o *repartiție gama-2 de parametrii* $\psi > 0$ *și* $\nu > 0$ dacă densitatea sa de probabilitate este

$$f_{\gamma 2}(x \mid \psi, \nu) = \begin{cases} \dfrac{e^{-\frac{\psi\nu x}{2}} x^{\frac{\nu}{2}-1}}{\left(\frac{\psi\nu}{2}\right)^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} & \text{pentru } x > 0; \\ 0 & \text{pentru } x \leqslant 0. \end{cases} \tag{13.2.3}$$

Mulțimea variabilelor aleatoare cu această repartiție va fi notată prin $\gamma 2(\psi, \nu)$. Este clar că $\tilde{x} \in \gamma 2(\psi, \nu)$ dacă și numai dacă $\left(\frac{\psi\nu}{2}\right)\tilde{x} \in \gamma^*\left(\frac{\nu}{2}\right) = \gamma\left(\frac{\nu}{2}, 1\right)$ sau, echivalent, $\tilde{x} \in \gamma\left(\frac{\nu}{2}, \frac{\psi\nu}{2}\right)$.

Mai departe, vom spune că o variabilă aleatoare are o *repartiție hi-pătrat cu* $\nu > 0$ *grade de libertate* dacă densitatea sa de probabilitate este

$$f_{\chi^2}(x \mid \nu) = \begin{cases} \dfrac{e^{-\frac{x}{2}} \cdot x^{\frac{\nu}{2}-1}}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} & \text{pentru } x > 0, \\ 0 & \text{pentru } x \leqslant 0. \end{cases} \tag{13.2.4}$$

Mulțimea variabilelor aleatoare cu această repartiție va fi notată prin $\chi^2(\nu)$. Este ușor de văzut că $\tilde{x} \in \chi^2(\nu)$ dacă și numai dacă $\tilde{x} \in \gamma\left(\frac{\nu}{2}, \frac{1}{2}\right)$ sau, echivalent, $\tilde{x} \in \gamma 2\left(\frac{1}{\nu}, \nu\right)$.

Scanned with OKEN Scanner

Repartiția hi-pătrat, care a fost descoperită de Helmert în 1876 și a fost pusă în valoare trei decenii mai tîrziu de Pearson, joacă un rol important în statistica matematică, ceea ce justifică tratarea sa separată.

Repartiția hi-pătrat introdusă mai sus este numită și *repartiția hi-pătrat standard*. Există și o variantă nestandardizată a acestei repartiții. Mai precis, spunem că o variabilă aleatoare are *repartiție hi-pătrat cu parametrii* $\nu > 0$ *și* $\sigma^2 > 0$, dacă densitatea sa de probabilitate este de forma

$$f_{\chi^2}(x|\nu, \sigma^2) = \begin{cases} \dfrac{e^{-x/2\sigma^2} \cdot x^{\nu/2-1}}{(2\sigma^2)^{\nu/2}\Gamma(\nu/2)} & \text{pentru } x>0, \\ 0 & \text{pentru } x \leqslant 0. \end{cases} \tag{13.2.5}$$

Mulțimea variabilelor aleatoare cu această repartiție va fi notată prin $\chi^2(\nu, \sigma^2)$. Este clar că $\chi^2(\nu, \sigma^2) = \gamma(\nu/2, 1/2\sigma^2)$, $\chi^2(\nu, \sigma^2) = \gamma 2(1/\nu\sigma^2, \nu)$ și $\chi^2(\nu, 1) = \chi^2(\nu)$.

Un alt caz particular interesant de repartiție gama a fost întîlnit de Erlang în studiile sale privind apelurile abonaților unei centrale telefonice. Vom spune că o variabilă aleatoare are o *repartiție Erlang de parametrii* $r > 0$ *și* $\lambda > 0$ dacă densitatea sa de probabilitate este de forma

$$f_{Er}(x|r, \lambda) = \begin{cases} \dfrac{\lambda r e^{-\lambda r x}(\lambda r x)^{r-1}}{\Gamma(r)} & \text{pentru } x>0, \\ 0 & \text{pentru } x \leqslant 0. \end{cases} \tag{13.2.6}$$

Mulțimea variabilelor aleatoare cu această repartiție este notată prin $Er(r, \lambda)$. Este ușor de văzut că $Er(r, \lambda) = \gamma(r, \lambda r)$ și că $Er(1, \lambda) = E(\lambda)$.

O altă repartiție din familia de repartiții gama este *repartiția gama inversă*. Spunem că o variabilă aleatoare are o *repartiție gama inversă de parametrii* $\psi > 0$ *și* $\nu > 0$ dacă densitatea de probabilitate corespunzătoare este de forma

$$f_{\gamma i}(x|\psi, \nu) = \begin{cases} \dfrac{e^{-\psi\nu/2x} \cdot x^{\nu/2+1}}{\left(\psi\nu/2\right)^{-\nu/2}\Gamma(\nu/2)} & \text{pentru } x>0, \\ 0 & \text{pentru } x \leqslant 0. \end{cases} \tag{13.2.7}$$

Mulțimea variabilelor aleatoare cu această repartiție va fi notată prin $\gamma i(\psi, \nu)$. Este clar că $\tilde{x} \in \gamma i(\psi, \nu)$ dacă și numai dacă $1/\tilde{x} \in \gamma 2(\psi, \nu)$.

Scanned with OKEN Scanner

În sfîrșit, spunem că o variabilă aleatoare are *repartiție gama-2 inversă de parametrii* $\psi > 0$ *și* $\nu > 0$ dacă densitatea sa de probabilitate este

$$f_{\gamma 2i}(x \mid \psi, \nu) = \begin{cases} \dfrac{2e^{-\psi\nu/2x^2} \cdot x^{-\nu-1}}{(\psi\nu/2)^{-\nu/2}\Gamma(\nu/2)} & \text{pentru } x>0, \\ 0 & \text{pentru } x \leqslant 0. \end{cases} \tag{13.2.8}$$

Mulțimea variabilelor aleatoare cu această repartiție va fi notată prin $\gamma 2i(\psi, \nu)$. Este ușor de văzut că $\tilde{x} \in \gamma 2i(\psi, \nu)$ dacă și numai dacă $\tilde{x}^2 \in \gamma i(\psi, \nu)$ sau, echivalent, dacă și numai dacă $\tilde{x}^{-2} \in \gamma 2(\psi, \nu)$.

13.2.1. REPARTIȚIA GAMA

Vom studia acum în mod amănunțit repartiția gama de parametrii $r > 0$ și $\lambda > 0$. Funcția caracteristică a unei variabile aleatoare $\tilde{x} \in \gamma^*(r)$ este

$$\varphi_{\gamma^*}(t \mid r) = (\Gamma(r))^{-1} \int_0^\infty e^{itx} e^{-x} x^{r-1}\, dx =$$

$$= (\Gamma(r))^{-1} \int_0^\infty \left(\sum_{n=0}^\infty \frac{(itx)^n}{n!} e^{-x} x^{r-1} \right) dx =$$

$$= (\Gamma(r))^{-1} \sum_{n=0}^\infty \frac{(it)^n}{n!} \int_0^\infty e^{-x} x^{n+r-1}\, dx =$$

$$= (\Gamma(r))^{-1} \sum_{n=0}^\infty \frac{\Gamma(n+r)}{n!} (it)^n =$$

$$= 1 + \sum_{n=1}^\infty \frac{r(r+1)\ldots(r+n-1)}{n!} (it)^n = (1-it)^{-r}$$

și deci

$$\varphi_{\gamma^*}(t \mid r) = (1-it)^{-r}, \tag{13.2.9}$$

de unde rezultă că funcția caracteristică a unei variabile aleatoare $\widetilde{x} \in \gamma(r, \lambda)$ este

$$\varphi_{\gamma}(t \mid r, \lambda) = \varphi_{1/\lambda \cdot \lambda \widetilde{x}}(t) = \varphi_{\lambda x}(t/\lambda), \tag{13.2.10}$$

unde $\lambda \widetilde{x} \in \gamma^*(r)$, conform observației făcute mai sus. Din relațiile (13.2.9) și (13.2.10) rezultă atunci că

$$\varphi_{\gamma}(t|r, \lambda) = (1-it/\lambda)^{-r} = \frac{\lambda^r}{(\lambda - it)^r}. \tag{13.2.11}$$

Dacă variabilele aleatoare $\widetilde{x}_j$, $1 \leqslant j \leqslant s$, sînt independente și $\widetilde{x}_j \in \gamma(r_j, \lambda)$, atunci variabila aleatoare

$$\widetilde{x} = \widetilde{x}_1 + \ldots + \widetilde{x}_s$$

aparține clasei $\gamma(r, \lambda)$, unde

$$r = r_1 + \ldots + r_s.$$

Într-adevăr, funcția caracteristică a lui $\widetilde{x}$ este dată de relația

$$\varphi_{\widetilde{x}}(t) = \varphi_{\widetilde{x}_1}(t) \ldots \varphi_{\widetilde{x}_s}(t) = \prod_{j=1}^{s} \left(\frac{\lambda}{\lambda - it}\right)^{r_j} = \left(\frac{\lambda}{\lambda - it}\right)^{r} = \varphi_{\gamma}(t \mid r, \lambda), \tag{13.2.12}$$

ceea ce demonstrează afirmația.

Momentele de ordinul s ale unei variabile aleatoare $\widetilde{x} \in \gamma(r, \lambda)$ se calculează direct:

$$M_s(\widetilde{x}) = (\Gamma(r))^{-1} \int_0^{\infty} x^s \lambda e^{-\lambda x} (\lambda x)^{r-1} \, dx = (\lambda^s \Gamma(r))^{-1} \int_0^{\infty} e^{-y} y^{r+s-1} \, dy$$

și deci

$$M_s(\widetilde{x}) = \frac{\Gamma(r+s)}{\lambda^s \Gamma(r)}. \tag{13.2.13}$$

Obținem deci

$$M(\tilde{x}) = r/\lambda; \quad M_2(\tilde{x}) = r(r+1)/\lambda^2; \quad M_3(\tilde{x}) = r(r+1)(r+2)/\lambda^3 \tag{13.2.14}$$

și apoi

$$V(\tilde{x}) = r/\lambda^2; \; V_3(\tilde{x}) = 2r/\lambda^3. \tag{13.2.15}$$

În sfîrșit, să observăm că dacă $\tilde{x} \in N(m, \sigma^2)$, atunci variabila aleatoare

$$\tilde{y} = \frac{(\tilde{x} - m)^2}{2\sigma^2}$$

aparține clasei $\gamma^*(1/2)$. Într-adevăr, pentru orice $x > 0$, avem

$$P(\tilde{y} < x) = P\left(-\sqrt{x} < \frac{\tilde{x} - m}{\sigma\sqrt{2}} < \sqrt{x}\right) = P\left(-\sqrt{2x} < \frac{\tilde{x} - m}{\sigma} < \sqrt{2x}\right).$$

Ținînd seama că variabila aleatoare $(\tilde{x} - m)/\sigma$ aparține clasei $N(0, 1)$, rezultă că putem scrie

$$P(\tilde{y} < x) = (2\pi)^{-1/2} \int_{-\sqrt{2x}}^{\sqrt{2x}} e^{-u^2/2}\, du = (2/\sqrt{2\pi}) \int_0^{\sqrt{2x}} e^{-u^2/2}\, du.$$

Dacă în ultima integrală se face schimbarea de variabilă $u^2 = 2v$ și se ține seama că $\Gamma(1/2) = \sqrt{\pi}$, rezultă ușor că

$$P(\tilde{y} < x) = (\Gamma(1/2))^{-1} \int_0^x e^{-v} v^{-1/2}\, dv,$$

ceea ce arată că densitatea de probabilitate a lui $\tilde{y}$ este

$$f_{\gamma^*}(x \mid 1/2) = \begin{cases} \dfrac{e^{-x} x^{1/2-1}}{\Gamma(1/2)} & \text{pentru } x > 0, \\ 0 & \text{pentru } x \leqslant 0, \end{cases}$$

ceea ce demonstrează afirmația.

13.2.2. REPARTIȚIA HI-PĂTRAT.

După cum am văzut mai sus, repartiția hi-pătrat cu parametrii $\nu > 0$ și $\sigma^2 > 0$ este repartiția gama cu parametrii $\nu/2$ și $1/2\sigma^2$, adică $\chi^2(\nu, \sigma^2) = \gamma(\nu/2, 1/2\sigma^2)$. De aici se deduc imediat următoarele rezultate, care se obțin prin particularizarea celor obținute în § 13.2.1 pentru repartiția gama.

Funcția caracteristică a unei variabile $\tilde{x} \in \chi^2(\nu, \sigma^2)$ este

$$\varphi_{\chi^2}(t \mid \nu, \sigma^2) = (1 - 2\sigma^2 ti)^{-\nu/2} \tag{13.2.16}$$

care se obține din (13.2.11) pentru $r = \nu/2$ și $\lambda = 1/2\sigma^2$.

Evident, dacă se consideră variabilele aleatoare independente $\tilde{x}_j \in \chi^2(\nu_j, \sigma^2)$, $1 \leqslant j \leqslant s$, se deduce ușor că suma lor este repartizată hi-pătrat cu parametrii ν și σ^2, unde

$$\nu = \nu_1 + \ldots + \nu_s,$$

cum rezultă imediat din cazul general.

Momentele de ordinul s ale unei variabile $\tilde{x} \in \chi^2(\nu, \sigma^2)$ sînt

$$M_s(\tilde{x}) = \frac{(2\sigma^2)^s \Gamma(s + \nu/2)}{\Gamma(\nu/2)} \tag{13.2.17}$$

și deci obținem

$$\begin{aligned} &M(\tilde{x}) = \nu\sigma^2; \quad M_2(\tilde{x}) = \nu(\nu + 2)\sigma^4; \\ &M_3(\tilde{x}) = \nu(\nu + 2)(\nu + 4)\sigma^6 \end{aligned} \tag{13.2.18}$$

și

$$V(\tilde{x}) = 2\nu\sigma^4; \; V_3(\tilde{x}) = 8\nu\sigma^6. \tag{13.2.19}$$

Teoremele care urmează arată legătura care se stabilește între repartiția hi-pătrat și celelalte repartiții.

Teorema 13.2.1. *Dacă variabilele aleatoare independente* $\tilde{x}_1, \ldots, \tilde{x}_\nu$ *sînt variabile aleatoare normale din clasa* $N(0, \sigma^2)$, *atunci variabila aleatoare*

$$\tilde{y} = \tilde{x}_1^2 + \ldots + \tilde{x}_\nu^2 \tag{13.2.20}$$

aparține clasei $\chi^2(\nu, \sigma^2)$.

Demonstrație. Deoarece $\tilde{x}_j \in N(0, \sigma^2)$, $1 \leqslant j \leqslant \nu$, din § 13.2.1 rezultă că variabilele aleatoare $\tilde{y}_j = \tilde{x}_j^2 / 2\sigma^2$ aparțin clasei $\gamma^*(1/2)$, ceea ce este echivalent cu faptul că $\tilde{x}_j^2 \in \gamma(1/2, 1/2\sigma^2)$ sau $\tilde{x}_j^2 \in \chi^2(1, \sigma^2)$. În sfîrșit, variabilele aleatoare $\tilde{x}_1^2, \ldots, \tilde{x}_\nu^2$ fiind independente, urmează că suma lor aparține clasei $\chi^2(\nu, \sigma^2)$, adică tocmai ceea ce trebuie demonstrat.

Teorema 13.2.2. *Fie* $\tilde{x}_1, \ldots, \tilde{x}_n$ *variabile aleatoare independente din clasa* $N(0, 1)$ *și fie*

$$\tilde{y}_i = \sum_{j=1}^{n} a_{ij} \tilde{x}_j, \quad 1 \leqslant i \leqslant \nu, \tag{13.2.21}$$

unde a_{ij} *sînt numere reale date, iar* $\nu < n$.

Dacă $\tilde{y}_1, \ldots, \tilde{y}_\nu$ *sînt variabile aleatoare independente din clasa* $N(0, 1)$, *atunci variabila aleatoare*

$$\tilde{z} = \sum_{j=1}^{n} \tilde{x}_j^2 - \sum_{j=1}^{\nu} \tilde{y}_j^2 \tag{13.2.22}$$

aparține clasei $\chi^2(n-\nu)$.

Demonstrație. Din faptul că $\tilde{y}_1, \ldots, \tilde{y}_\nu$ sînt independente urmează că

$$M(\tilde{y}_i \tilde{y}_j) = M(\tilde{y}_i) M(\tilde{y}_j) = \delta_{ij}, \quad 1 \leqslant i, j \leqslant \nu,$$

unde δ_{ij} este simbolul lui Krönecker.

Analog, deoarece $\tilde{x}_1, \ldots, \tilde{x}_n$ sînt independente, obținem

$$M(\tilde{x}_i \tilde{x}_j) = \delta_{ij}, \quad 1 \leqslant i, j \leqslant n.$$

Utilizînd aceste relații obținem:

$$M(\tilde{y}_i \tilde{y}_j) = M\left[\left(\sum_{k=1}^{n} a_{ik} \tilde{x}_k\right)\left(\sum_{s=1}^{n} a_{js} \tilde{x}_s\right)\right] =$$

$$= M\left(\sum_{k,s=1}^{n} a_{ik} a_{js} \tilde{x}_k \tilde{x}_s\right) = \sum_{k,s=1}^{n} a_{ik} a_{js} M(\tilde{x}_k \tilde{x}_s) =$$

$$= \sum_{k,s=1}^{n} a_{ik} a_{js} \delta_{ks} = \sum_{k=1}^{n} a_{ik} a_{jk} = \delta_{ij}.$$

Scanned with OKEN Scanner

ceea ce arată că coeficienții a_{ij}, $1 \leqslant i \leqslant \nu$, $1 \leqslant j \leqslant n$, satisfac condiții de ortogonalitate. Rezultă deci că putem completa matricea $(a_{ij})_{\substack{1 \leqslant i \leqslant \nu \\ 1 \leqslant j \leqslant n}}$ pînă la o matrice pătrată de ordinul n care să verifice condiția de ortogonalitate, introducînd în acest scop coeficienții a_{ij}, $\nu+1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant n$.

Prin urmare vom avea

$$\sum_{k=1}^{n} a_{ik}a_{jk} = \sum_{k=1}^{n} a_{ki}a_{kj} = \delta_{ij}, \quad 1 \leqslant i, \ j \leqslant n.$$

Dacă introducem variabilele aleatoare

$$\tilde{y}_i = \sum_{k=1}^{n} a_{ik}\tilde{x}_k, \quad \nu+1 \leqslant i \leqslant n,$$

se poate vedea ușor că $\tilde{y}_1, \ldots, \tilde{y}_n$ sînt variabile aleatoare normale necorelate și deci independente. Este clar că

$$\sum_{i=1}^{n} \tilde{y}_i^2 = \sum_{i=1}^{n} \left(\sum_{j,\,k=1}^{n} a_{ij}a_{ik}\tilde{x}_j\tilde{x}_k \right) =$$

$$= \sum_{j,\,k=1}^{n} \left(\sum_{i=1}^{n} a_{ij}a_{ik} \right) \tilde{x}_j\tilde{x}_k = \sum_{j,\,k=1}^{n} \delta_{jk}\tilde{x}_j\tilde{x}_k = \sum_{j=1}^{n} \tilde{x}_j^2$$

și deci

$$\tilde{z} = \sum_{j=1}^{n} \tilde{x}_j^2 - \sum_{j=1}^{n} \tilde{y}_j^2 + \sum_{j=\nu+1}^{n} \tilde{y}_j^2 = \sum_{j=\nu+1}^{n} \tilde{y}_j^2.$$

Cum $\tilde{y}_{\nu+1}, \ldots, \tilde{y}_n$ sînt variabile aleatoare independente din clasa $N(0, 1)$, aplicînd teorema precedentă rezultă că $\tilde{z} \in \chi^2(n-\nu, 1) = \chi^2(n-\nu)$, ceea ce dealtfel trebuia demonstrat.

Teorema 13.2.3. *Dacă* $\tilde{\mathbf{x}}$ *este un vector aleator normal nesingular din clasa* $\mathrm{N}(\mathbf{m}, \boldsymbol{\delta}^2)$, *atunci variabila aleatoare*

$$\tilde{y} = (\tilde{\mathbf{x}} - \mathbf{m})'\boldsymbol{\delta}^{-2}(\tilde{\mathbf{x}} - \mathbf{m})$$

aparține clasei $\chi^2(\mathrm{k})$, *unde* k *este numărul componentelor lui* $\tilde{\mathbf{x}}$.

Scanned with OKEN Scanner

Demonstrație. Deoarece $\boldsymbol{\delta}^2$ este pozitiv definită, putem determina o matrice $\mathbf{D}$ astfel încît $\mathbf{D}\boldsymbol{\delta}^2\mathbf{D}' = \mathbf{I}$. Vectorul aleator

$$\tilde{\mathbf{z}} = \mathbf{D}(\tilde{\mathbf{x}} - \mathbf{m})$$

este un vector normal cu media nulă, iar matricea sa de covarianță este

$$V(\tilde{\mathbf{z}}) = M[(\mathbf{D}(\tilde{\mathbf{x}} - \mathbf{m}))(\mathbf{D}(\tilde{\mathbf{x}} - \mathbf{m}))'] =$$

$$= M(\mathbf{D}(\tilde{\mathbf{x}} - \mathbf{m})(\tilde{\mathbf{x}} - \mathbf{m})'\mathbf{D}') = \mathbf{D}\boldsymbol{\delta}^2\mathbf{D}' = \mathbf{I}.$$

Urmează că $\tilde{\mathbf{z}}$ aparține clasei $N(\mathbf{0}, \mathbf{I})$ și deci componentele sale sînt variabile aleatoare normale standard independente.

Pe de altă parte, avem

$$\tilde{z}_1^2 + \ldots + \tilde{z}_k^2 = \tilde{\mathbf{z}}'\tilde{\mathbf{z}} = (\mathbf{D}(\tilde{\mathbf{x}} - \mathbf{m}))'(\mathbf{D}(\tilde{\mathbf{x}} - \mathbf{m})) =$$

$$= (\tilde{\mathbf{x}} - \mathbf{m})'\mathbf{D}'\mathbf{D}(\tilde{\mathbf{x}} - \mathbf{m}) = (\tilde{\mathbf{x}} - \mathbf{m})'\boldsymbol{\delta}^{-2}(\tilde{\mathbf{x}} - \mathbf{m}) = \tilde{y},$$

de unde reiese că $\tilde{y} \in \chi^2(k)$, conform teoremei (13.2.1).

Amintim de asemenea faptul că repartiția hi-pătrat este o repartiție limită. Astfel, dacă vectorul aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_k)'$ are repartiție multinomială cu parametrii n și $\mathbf{p} = (p_1, \ldots, p_k)'$, atunci, notînd ca de obicei

$$p_0 = 1 - (p_1 + \ldots + p_k),$$

$$\tilde{x}_0 = 1 - (\tilde{x}_1 + \ldots + \tilde{x}_k),$$

rezultă că variabila aleatoare

$$\tilde{y} = \sum_{j=0}^{k} (1/np_j)(\tilde{x} - np_j)^2 \qquad (13.2.23)$$

este asimptotic (pentru $n \to +\infty$) repartizată hi-pătrat cu k grade de libertate. Statistica (13.2.23) este numită *statistica hi-pătrat a lui Pearson* și joacă un rol important în rezolvarea anumitor probleme de statistică matematică.

Scanned with OKEN Scanner

13.3. FAMILIA DE REPARTIȚII BETA

13.3.1. REPARTIȚIA BETA (STANDARD)

Spunem că o variabilă aleatoare are *repartiție beta (standard) de parametrii* $\rho > 0$ *și* $\nu > 0$ dacă densitatea sa de probabilitate este

$$f_\beta(x \mid \rho, \nu) = \begin{cases} \dfrac{x^{\rho-1}(1-x)^{\nu-1}}{B(\rho, \nu)} & \text{pentru } x \in (0, 1), \\ 0 & \text{în rest,} \end{cases} \tag{13.3.1}$$

unde $B(\rho, \nu)$ este funcția beta* definită pentru orice $\rho > 0$ și $\nu > 0$ prin relația

$$B(\rho, \nu) = \int_0^1 x^{\rho-1}(1-x)^{\nu-1}\,dx. \tag{13.3.2}$$

Mulțimea variabilelor aleatoare cu această repartiție va fi notată prin $\beta(\rho, \nu)$. Este clar că $\tilde{x} \in \beta(\rho, \nu)$ dacă și numai dacă $(1 - \tilde{x}) \in \beta(\nu, \rho)$.

Momentul de ordinul r al unei variabile aleatoare $\tilde{x} \in \beta(\rho, \nu)$ se calculează direct:

$$M_r(\tilde{x}) = (B(\rho, \nu))^{-1} \int_0^1 x^{r+\rho-1}(1-x)^{\nu-1}\,dx = \frac{B(r+\rho, \nu)}{B(\rho, \nu)}. \tag{13.3.3}$$

Dacă r este întreg, din (13.3.3) se obține ușor (vezi Anexa, relația (A.2.9)):

$$M_r(\tilde{x}) = \frac{\rho(\rho+1)\ldots(\rho+r-1)}{(\rho+\nu)(\rho+\nu+1)\ldots(\rho+\nu+r+1)} \tag{13.3.4}$$

și, în particular,

$$M(\tilde{x}) = \frac{\rho}{\rho+\nu}, \quad M_2(\tilde{x}) = \frac{\rho(\rho+1)}{(\rho+\nu)(\rho+\nu+1)}. \tag{13.3.5}$$

* Vezi Anexa, § A.2.

Rezultă că dispersia unei variabile aleatoare $\tilde{x} \in \beta(\rho, \nu)$ este

$$V(\tilde{x}) = \frac{\rho\nu}{(\rho+\nu)^2(\rho+\nu+1)}. \tag{13.3.6}$$

Este de asemenea interesant rezultatul următor, care face precizări în legătură cu raportul unor variabile aleatoare cu repartiție gama.

Teorema 13.3.1. *Fie* $\tilde{x}$ *și* $\tilde{y}$ *două variabile aleatoare independente, unde* $\tilde{x} \in \gamma^*(\rho)$ *și* $\tilde{y} \in \gamma^*(\nu)$. *În aceste condiții, variabila aleatoare*

$$\tilde{z} = \frac{\tilde{x}}{\tilde{x}+\tilde{y}} \tag{13.3.7}$$

aparține clasei $\beta(\rho, \nu)$.

Demonstrație. Deoarece x și $\tilde{y}$ sînt variabile aleatoare independente, rezultă că densitatea de probabilitate a vectorului aleator $(\tilde{x}, \tilde{y})$ este

$$f(x, y) = f_{\gamma^*}(x \mid \rho) f_{\gamma^*}(y \mid \nu). \tag{13.3.8}$$

Dacă introducem variabila aleatoare

$$\tilde{z}' = \tilde{x} + \tilde{y},$$

rezultă ușor că putem scrie

$$\tilde{x} = \tilde{z}\tilde{z}', \quad \tilde{y} = \tilde{z}'(1 - \tilde{z}).$$

Densitatea de probabilitate $f(z', z)$ a vectorului aleator $(\tilde{z}', \tilde{z})'$ se obține pe calea obișnuită, ținînd seama că

$$\frac{D(x, y)}{D(z', z)} = \begin{vmatrix} z & z' \\ 1-z & -z' \end{vmatrix} = -z'.$$

Obținem

$$f(z', z) = f(x, y)z'.$$

Cu alte cuvinte, ținînd seama și de relația (13.3.8), deducem

$$f(z', z) = \frac{1}{\Gamma(\rho)\Gamma(\nu)} e^{-z'}(z')^{\rho+\nu-1} z^{\rho-1}(1-z)^{\nu-1}.$$

Prin urmare densitatea de probabilitate a lui $\tilde{z}$ este

$$f(z) = \int_0^\infty f(z',\ z)\, dz' = \frac{\Gamma(\rho + \nu)}{\Gamma(\rho)\,\Gamma(\nu)} z^{\rho-1}(1-z)^{\nu-1}$$

sau

$$f(z) = \frac{z^{\rho-1}(1-z)^{\nu-1}}{B(\rho,\ \nu)},$$

ceea ce arată că $\tilde{z} \in \beta(\rho, \nu)$.

Din această teoremă obținem imediat rezultatul următor.

Teorema 13.3.2. *Fie* $\tilde{x}_1, \ldots, \tilde{x}_n$ *variabile aleatoare independente din clasa* $N(0, \sigma^2)$. *Dacă* $1 \leqslant \nu < n$, *atunci variabila aleatoare*

$$\tilde{y} = \frac{\tilde{x}_1^2 + \ldots + \tilde{x}_\nu^2}{\tilde{x}_1^2 + \ldots + \tilde{x}_n^2}$$

aparține clasei $\beta(\nu/2,\ (n-\nu)/2)$.

Demonstrație. Fie

$$\tilde{y}_1 = \tilde{x}_1^2 + \ldots + \tilde{x}^2,$$

$$\tilde{y}_2 = \tilde{x}_{\nu+1}^2 + \ldots + \tilde{x}_n^2.$$

Din teorema 13.2.1 rezultă că $\tilde{y}_1 \in \chi^2(\nu, \sigma^2)$ și $\tilde{y}_2 \in \chi^2(n-\nu, \sigma^2)$, ceea ce este echivalent cu $\tilde{y}_1 \in \gamma(\nu/2, 1/2\sigma^2)$ și $\tilde{y}_2 \in \gamma((n-\nu)/2, 1/2\sigma^2)$. Din ultimele două apartenențe rezultă de asemenea că $(1/2\sigma^2)\tilde{y}_1 \in \gamma^*(\nu/2)$ și $(1/2\sigma^2)\tilde{y}_2 \in \gamma^*((n-\nu)/2)$. Cum $\tilde{y}_1$ și $\tilde{y}_2$ sînt independente, din teorema precedentă rezultă că

$$\tilde{y} = \frac{\frac{1}{2\sigma^2}\tilde{y}_1}{\frac{1}{2\sigma^2}\tilde{y}_1 + \frac{1}{2\sigma^2}\tilde{y}_2}$$

aparține clasei $\beta(\nu/2,\ (n-\nu)/2)$, ceea ce trebuia demonstrat.

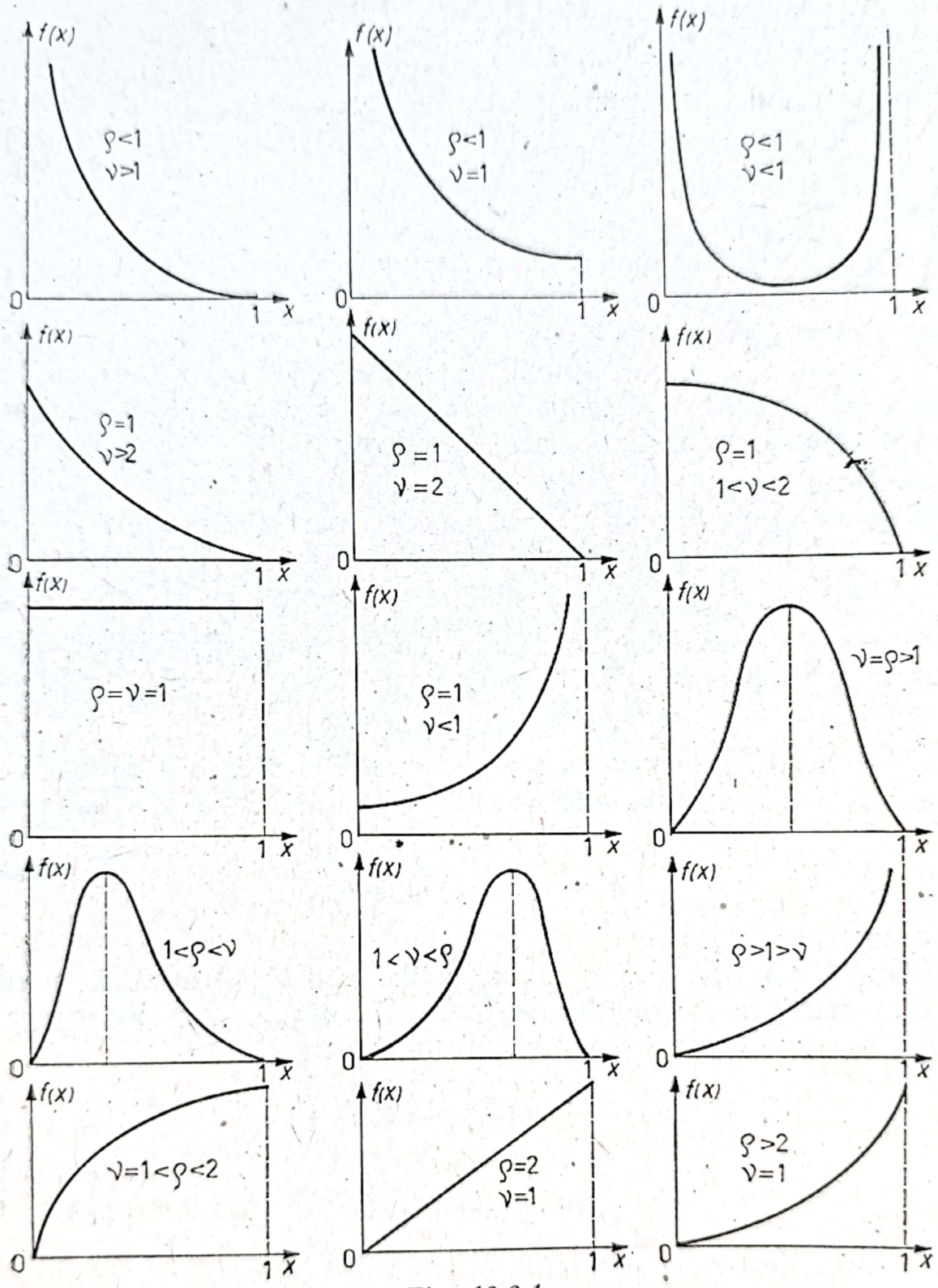


Fig. *13.3.1*

În figura 13.3.1 sînt date cîteva exemple de grafice ale densității de probabilitate ale unei repartiții beta pentru cîteva valori remarcabile ale parametrilor ρ și ν. Dintr-o simplă examinare a acestei figuri se poate vedea marea varietate de

forme pe care le poate avea graficul densității de probabilitate a repartiției beta. Aceasta explică în bună măsură importanța acestei repartiții ca instrument cu ajutorul căruia putem exprima opinii privind o proporție necunoscută $p \in (0, 1)$.

13.3.2. REPARTIȚIA BETA NESTANDARDIZATĂ.

Se spune că o variabială aleatoare are o *repartiție beta (nestandardizată) de parametrii* $\rho > 0$, $\nu > 0$, $a \in R$ *și* $b > a$ dacă densitatea sa de probabilitate este de forma

$$f_\beta(x|\rho, \nu, a, b) = \begin{cases} \dfrac{(x-a)^{\rho-1}(b-x)^{\nu-1}}{(b-a)^{\rho+\nu-1}B(\rho, \nu)} & \text{pentru } x \in (a, b), \\ 0 & \text{pentru } x \notin (a, b). \end{cases} \tag{13.3.9}$$

Mulțimea variabilelor aleatoare cu această repartiție va fi notată prin $\beta(\rho, \nu, a, b)$. Este clar că

$$\beta(\rho, \nu, 0, 1) = \beta(\rho, \nu).$$

Mai mult, se poate vedea ușor că $\tilde{x} \in \beta(\rho, \nu, a, b)$ dacă și numai dacă variabila aleatoare

$$\tilde{y} = \frac{\tilde{x} - a}{b - a} \tag{13.3.10}$$

aparține clasei $\beta(\rho, \nu)$.

Din (13.3.5), (13.3.6) și (13.3.10) rezultă atunci că media și dispersia unei variabile aleatoare $\tilde{x} \in \beta(\rho, \nu, a, b)$ pot fi obținute în modul următor. Avem mai întîi

$$M(\tilde{x}) = a + (b-a)M(\tilde{y}) = a + (b-a)\frac{\rho}{\rho+\nu}$$

sau

$$M(\tilde{x}) = \frac{a\nu + b\rho}{\nu + \rho} \tag{13.3.11}$$

și apoi

$$V(\tilde{x}) = (b-a)^2 V(\tilde{y}),$$

de unde

$$V(\tilde{x}) = \frac{(b-a)^2 \rho\nu}{(\rho+\nu)^2(\rho+\nu+1)}. \tag{13.3.12}$$

Este de asemenea util să observăm că funcția de repartiție a unei variabile aleatoare $\tilde{x} \in \beta(\rho, \nu, a, b)$ se poate exprima astfel:

$$F_\beta(x \mid \rho, \nu, a, b) = P(\tilde{x} < x) = P(a + (b-a)\tilde{y} < x) =$$

$$= P\left(\tilde{y} < \frac{x-a}{b-a}\right) = F_\beta\left(\frac{x-a}{b-a} \,\middle|\, \rho, \nu\right),$$

unde $F_\beta(y \mid \rho, \nu)$ este funcția de repartiție a unei variabile aleatoare $\tilde{y} \in \beta(\rho, \nu)$.

Repartiția beta nestandardizată se utilizează uneori pentru a exprima opinii privind durata anumitor lucrări parțiale ale unui proiect mai mare. Astfel de probleme apar în așa-numita *metodă a drumului critic* sau *metoda PERT* (vezi, de exemplu, A. Kaufmann și G. Desbazeille (1964)).

Modul m al repartiției beta nestandardizate se determină imediat ca rădăcină a ecuației

$$\frac{\mathrm{d}}{\mathrm{d}x}(f_\beta(x \mid \rho, \nu, a, b)) = 0$$

și se obține

$$m = \frac{a(\nu-1) + b(\rho-1)}{\rho + \nu - 2} \tag{13.3.13}$$

Rezultă atunci că media unei variabile aleatoare $\tilde{x} \in \beta(\rho, \nu, a, b)$ poate fi scrisă sub forma

$$M(\tilde{x}) = \frac{a + b + m(\rho + \nu - 2)}{\rho + \nu}. \tag{13.3.14}$$

În metoda PERT se adoptă, de regulă, următoarele valori ale parametrilor ρ și ν:

$$\rho = 3 + \sqrt{2}; \quad \nu = 3 - \sqrt{2} \tag{13.3.15}$$

sau

$$\rho = 3 - \sqrt{2}; \quad \nu = 3 + \sqrt{2}, \tag{13.3.16}$$

Scanned with OKEN Scanner

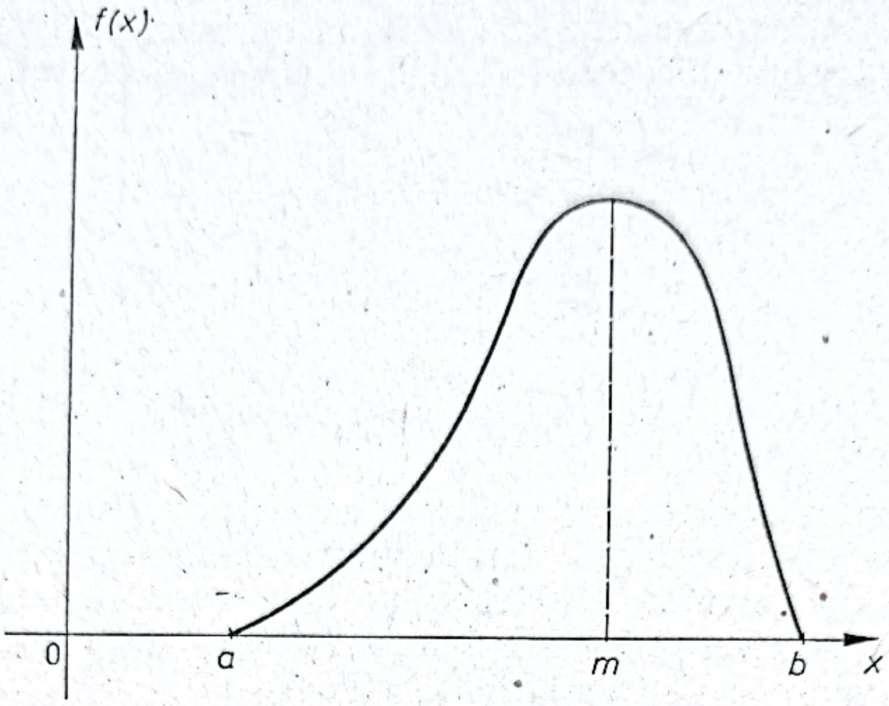


Fig. 13.3.2

caz în care $\rho + \nu = 6$. Rezultă atunci că media și dispersia se pot scrie în acest caz particular astfel:

$$M(\tilde{x}) = \frac{a + b + 4m}{6} \tag{13.3.17}$$

și

$$V(\tilde{x}) = \left(\frac{b - a}{6}\right)^2 . \tag{13.3.18}$$

Cu ajutorul formulelor (13.3.17) și (13.3.18), estimarea mediei și dispersiei duratei aleatoare $\tilde{x}$ a lucrării se face utilizînd o *durată minimă* a, o *durată maximă* b și o *durată cea mai probabilă* m.

În figura 13.3.2 este prezentată forma graficului densității beta nestandardizată pentru valorile (13.3.15) ale parametrilor ρ și ν.

Un caz particular important al repartiției beta este repartiția *uniformă* (sau *rectangulară*). O variabilă aleatoare are repartiție uniformă de parametrii $a \in R$ și $b > a$ dacă densitatea

sa de probabilitate este de forma

$$f_u(x \mid a, b) = \begin{cases} \dfrac{1}{b-a} & \text{pentru } x \in [a, b], \\ 0 & \text{pentru } x \notin [a, b]. \end{cases}$$

Mulțimea variabilelor aleatoare cu această repartiție este notată prin $U(a, b)$. Se vede ușor că $U(a, b) = \beta(1, 1, a, b)$ și deci media și dispersia acestei repartiții sînt date de formulele:

$$M(\tilde{x}) = \frac{a+b}{2}, \quad V(\tilde{x}) = \frac{(b-a)^2}{12}.$$

Dealtfel, un calcul direct ne conduce la concluzia că momentul de ordinul r al unei variabile aleatoare $\tilde{x} \in U(a, b)$ este

$$M_r(\tilde{x}) = \frac{1}{r+1} \cdot \frac{b^{r+1} - a^{r+1}}{b-a}.$$

Se vede de asemenea ușor că media $M(\tilde{x}) = m$ și dispersia $V(\tilde{x}) = \sigma^2$ ale unei repartiții oarecare coincid cu media și dispersia repartiției uniforme de parametrii $a = m - \sigma\sqrt{3}$, $b = m + \sigma\sqrt{3}$. Acest fapt ne permite să utilizăm repartiția uniformă ca o măsură a împrăștierii valorilor unei repartiții oarecare.

13.3.3. REPARTIȚIA BETA-2 INVERSĂ

Spunem că o variabilă aleatoare $\tilde{x}$ are o *repartiție beta-2 inversă cu parametrii* $\rho > 0$, $\nu > 0$, *și* $b > 0$ dacă densitatea sa de probabilitate este

$$f_{\beta 2i}(x \mid \rho, \nu, b) = \begin{cases} \dfrac{x^{\rho-1} b^{\nu}}{B(\rho, \nu)(b+x)^{\rho+\nu}} & \text{pentru } x > 0, \\ 0 & \text{pentru } x \leqslant 0. \end{cases} \tag{13.3.19}$$

Mulțimea variabilelor aleatoare cu această repartiție va fi notată prin $\beta 2i(\rho, \nu, b)$. Se poate verifica fără dificultate

că $\tilde{x} \in \beta 2i(\rho, \nu, b)$ dacă și numai dacă variabila aleatoare

$$\tilde{y} = \frac{\tilde{x}}{\tilde{x} + b} \tag{13.3.20}$$

aparține clasei $\beta(\rho, \nu)$. Funcția de repartiție a unei variabile aleatoare $\tilde{x} \in \beta\, 2i(\rho, \nu, b)$ se poate exprima în modul următor

$$F_{\beta 2i}(x \mid \rho, \nu, b) = P(\tilde{x} < x) = P\left(\frac{b\tilde{y}}{1 - \tilde{y}} < x\right) =$$

$$= P\left(\tilde{y} < \frac{x}{b + x}\right) = F_{\beta}\left(\frac{x}{b + x} \,\middle|\, \rho, \nu\right). \tag{13.3.21}$$

Momentul de ordinul r al unei variabile aleatoare $\tilde{x} \in \beta 2i(\rho, \nu, b)$ se obține prin calcul direct astfel:

$$M_r(\tilde{x}) = \int_0^{\infty} x^r f_{\beta 2i}(x \mid \rho, \nu, b)\, dx = \frac{b}{B(\rho, \nu)} \int_0^{\infty} x^{r+\rho-1}(b + x)^{-(\rho+\nu)}\, dx.$$

Făcînd în ultima integrală schimbarea de variabilă $x(b+x)^{-1} = y$ se obține

$$M_r(\tilde{x}) = \frac{b^r \Gamma(\rho + r)\Gamma(\nu - r)}{\Gamma(\rho)\Gamma(\nu)}, \tag{13.3.22}$$

cu condiția ca $0 < r < \nu$. Avem atunci

$$M(\tilde{x}) = \frac{b\rho}{\nu - 1} \tag{13.3.23}$$

dacă $\nu > 1$ și

$$M_2(\tilde{x}) = \frac{b^2 \rho(\rho + 1)}{(\nu - 1)(\nu - 2)}, \tag{13.3.24}$$

cu condiția ca $\nu > 2$.

Urmează atunci că dispersia este

$$V(\tilde{x}) = \frac{b^2 \rho(\rho + \nu - 1)}{(\nu - 1)^2(\nu - 2)}, \tag{13.3.25}$$

cu condiția ca $\nu > 2$.

Scanned with OKEN Scanner

13.3.4. REPARTIȚIA F

Un caz particular important al repartiției beta-2 inversă este *repartiția F*. Vom spune că o variabilă aleatoare are *repartiție F cu* $\nu_1 > 0$ *și* $\nu_2 > 0$ *grade de libertate* * dacă densitatea sa de probabilitate este

$$f_F(x \mid \nu_1, \nu_2) = \begin{cases} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{\nu_2}} \frac{x^{\frac{\nu_1}{2}-1}}{B\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}\right)\left(1+\frac{\nu_1}{\nu_2}x\right)^{\frac{\nu_1+\nu_2}{2}}} & \text{pentru } x>0, \\ 0 & \text{pentru } x \leqslant 0. \end{cases} \tag{13.3.26}$$

Mulțimea variabilelor aleatoare cu această repartiție va fi notată prin $F(\nu_1, \nu_2)$. Este ușor de văzut că

$$F(\nu_1, \nu_2) = \beta 2i(\nu_1/2,\ \nu_2/2,\ \nu_2/\nu_1), \tag{13.3.27}$$

de unde rezultă imediat că momentele de ordinul r ale unei variabile aleatoare $\tilde{x} \in F(\nu_1, \nu_2)$ sînt

$$M_r(\tilde{x}) = \left(\frac{\nu_2}{\nu_1}\right)^r \frac{\Gamma\left(\frac{\nu_1}{2}+r\right)\Gamma\left(\frac{\nu_2}{2}-r\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}, \tag{13.3.28}$$

cu condiția ca $0 < r < \nu_2/2$. Pentru r întreg, se obține

$$M_r(\tilde{x}) = \left(\frac{\nu_2}{\nu_1}\right)^r \frac{\nu_1(\nu_1+2)\ldots(\nu_1+2r-2)}{(\nu_2-2)(\nu_2-4)\ldots(\nu_2-2r)}, \tag{13.3.29}$$

cu condiția ca $0 < r < \nu_2/2$. În particular, avem:

$$M(\tilde{x}) = \frac{\nu_2}{\nu_2-2};\ M_2(\tilde{x}) = \frac{\nu_2^2}{\nu_1} \cdot \frac{\nu_1+2}{(\nu_2-2)(\nu_2-4)}, \tag{13.3.30}$$

valabile respectiv pentru $\nu_2 > 2$ și $\nu_2 > 4$.

---

* Repartiția F mai este numită și *repartiția Snedecor*.

Scanned with OKEN Scanner

Dispersia variabilei aleatoare $\tilde{x} \in F(\nu_1, \nu_2)$ este deci

$$V(\tilde{x}) = \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{(\nu_2 - 2)}, \qquad (13.3.31)$$

cu condiția ca $\nu_2 > 4$. Este de remarcat că media unei variabile aleatoare din clasa $F(\nu_1, \nu_2)$ nu depinde de ν_1.

Legătura dintre repartiția F și celelalte repartiții este precizată în teoremele care urmează.

Teorema 13.3.3. *Dacă* $\tilde{x}_1 \in \chi^2(\nu_1, \sigma^2)$ *și* $\tilde{x}_2 \in \chi^2(\nu_2, \sigma^2)$ *sînt variabile aleatoare independente, atunci variabila aleatoare*

$$y = \frac{\frac{1}{\nu_1}\tilde{x}_1}{\frac{1}{\nu_2}\tilde{x}_2}$$

aparține clasei $F(\nu_1, \nu_2)$.

Demonstrație. Deoarece $\tilde{x}_1$ și $\tilde{x}_2$ sînt variabile aleatoare independente, rezultă că densitatea de probabilitate $f(x_1, x_2)$ a vectorului aleator $(\tilde{x}_1, \tilde{x}_2)'$ este dată de relația

$$f(x_1, x_2) = f_{\chi^2}(x_1 \mid \nu_1, \sigma^2)\, f_{\chi^2}(x_2 \mid \nu_2, \sigma^2).$$

Densitatea de probabilitate a vectorului aleator $(\tilde{y}, \tilde{x}_2)'$ se obține imediat, observînd că

$$x_1 = \frac{\nu_1}{\nu_2} y x_2; \quad x_2 = x_2$$

și deci

$$\frac{D(x_1, x_2)}{D(y, x_2)} = \begin{vmatrix} \frac{\nu_1}{\nu_2} x_2 & \frac{\nu_1}{\nu_2} y \\ 0 & 1 \end{vmatrix} = \frac{\nu_1}{\nu_2} x_2,$$

de unde rezultă

$$f(y, x_2) = f(x_1, x_2) \frac{\nu_1}{\nu_2} x_2 =$$

$$= f_{\chi^2}\left(\frac{\nu_1}{\nu_2} y x_2 \mid \nu_1, \sigma^2\right) f_{\chi^2}(x_2 \mid \nu_2, \sigma^2) \frac{\nu_1}{\nu_2} x_2.$$

Scanned with OKEN Scanner

Rezultă atunci că densitatea de probabilitate a lui $\tilde{y}$ este

$$f(y)=\int_{-\infty}^{\infty} f(y, x_2)\,\mathrm{d}x_2 = \frac{\nu_1}{\nu_2}\int_0^{\infty} f_{\chi^2}\left(\frac{\nu_1}{\nu_2}\, yx_2 \mid \nu_1, \sigma^2\right) f_{\chi^2}(x_2 \mid \nu_2, \sigma^2)\, x_2\,\mathrm{d}x_2.$$

După calcule elementare se obține $f(y) = f_F(y \mid \nu_1, \nu_2)$ și teorema este demonstrată.

Această teoremă, împreună cu teorema (13.2.1), conduce la consecința următoare.

Teorema 13.3.4. *Dacă* $\tilde{x}_1, \ldots, \tilde{x}_{\nu_1}$ *și* $\tilde{y}_1, \ldots, \tilde{y}_{\nu_2}$ *sînt variabile aleatoare independente din clasa* $N(0, \sigma^2)$, *atunci variabila aleatoare*

$$\tilde{z} = \frac{\nu_2}{\nu_1} \cdot \frac{\tilde{x}_1^2 + \ldots + \tilde{x}_{\nu_1}^2}{\tilde{y}_1^2 + \ldots + \tilde{y}_{\nu_2}^2}$$

aparține clasei $F(\nu_1, \nu_2)$.

13.4. REPARTIȚIA STUDENT

O altă repartiție importantă în problema de inferență statistică este *repartiția Student*. Spunem că o variabilă aleatoare are *repartiție Student (standard) cu* $\nu > 0$ *grade de libertate* dacă densitatea sa de probabilitate este

$$f_{S^*}(x \mid \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \qquad (13.4.1)$$

definită pentru orice $x \in R$. Mulțimea variabilelor aleatoare cu această repartiție este notată prin $S^*(\nu)$.

Deoarece $f_{S^*}(x \mid \nu)$ este o funcție pară, rezultă că media unei variabile aleatoare $\tilde{x} \in S^*(\nu)$ este nulă, iar momentele centrate coincid cu momentele obișnuite.

Momentele de ordin impar sînt

$$M_{2r+1}(\tilde{x}) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)} \int_{-\infty}^{\infty} x^{2r+1}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} dx = 0 \quad (13.4.2)$$

deoarece funcția de integrat este impară.

Pentru calculul momentelor de ordin par vom face substituția $x^2 = \nu y$ în integrala care apare mai jos și vom obține

$$M_{2r}(\tilde{x}) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)} \int_{-\infty}^{+\infty} x^{2r}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} dx =$$

$$= \frac{\nu^r \Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \int_0^{\infty} y^{r-\frac{1}{2}}(1+y)^{-\frac{\nu+1}{2}} dy.$$

Ținînd seama de formula (A.2.4) din Anexă, rezultă că ultima integrală are valoarea $B\left(\frac{\nu}{2}-r,\ r+1/2\right)$ și deci

$$M_{2r}(\tilde{x}) = \frac{\nu^r \Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \cdot \frac{\Gamma\left(\frac{\nu}{2}-r\right)\Gamma\left(r+\frac{1}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)}, \quad (13.4.3)$$

cu condiția ca $0 < r < \nu/2$.

Ținînd seama că pentru $r < \frac{\nu}{2}$ întreg putem scrie

$$\Gamma\left(\frac{\nu}{2}\right) = \left(\frac{\nu}{2}-1\right)\cdots\left(\frac{\nu}{2}-r\right)\Gamma\left(\frac{\nu}{2}-r\right)$$

Scanned with OKEN Scanner

și

$$\Gamma\left(r+\frac{1}{2}\right)=\left(r-\frac{1}{2}\right)\cdots\frac{1}{2}\,\Gamma\left(\frac{1}{2}\right),$$

rezultă că

$$M_{2r}(\tilde{x})=\frac{1\cdot 3\ldots(2r-1)}{(\nu-2)(\nu-4)\ldots(\nu-2r)}\cdot\nu^r, \qquad (13.4.4)$$

cu condiția ca $r<\frac{\nu}{2}$. Primele două momente de ordin par, în cazurile cînd există, sînt:

$$M_2(\tilde{x})=V(\tilde{x})=\frac{\nu}{\nu-2};\quad M_4(\tilde{x})=V_4(\tilde{x})=\frac{3\nu^2}{(\nu-2)(\nu-4)}. \qquad (13.4.5)$$

Din relația (13.4.4) rezultă imediat că

$$\lim_{\nu\to\infty} M_{2r}(\tilde{x})=1\cdot 3\cdot\ldots(2r-1).$$

Cum $M_{2r+1}(\tilde{x})=0$, rezultă că momentele repartiției $S^*(\nu)$ tind la momentele repartiției $N(0,1)$ cînd $\nu\to+\infty$. Aceasta este suficient pentru a justifica teorema care urmează.

Teorema 13.4.1. *Dacă* $\tilde{x}\in S^*(\nu)$, *atunci* $\tilde{x}$ *este asimptotic (pentru* $\nu\to+\infty$*) normală din clasa* $N(0,1)$.

Teoremele următoare arată legătura dintre repartiția Student și celelalte repartiții cunoscute.

Teorema 13.4.2. *Dacă* $\tilde{x}\in N(0,\sigma^2)$ *și* $\tilde{y}\in\chi^2(\nu,\sigma^2)$ *sînt variabile aleatoare independente, atunci variabila aleatoare*

$$\tilde{z}=\tilde{x}/\sqrt{\tilde{y}/\nu} \qquad (13.4.6)$$

aparține clasei $S^*(\nu)$.

Demonstrație. Densitatea de probabilitate $f(x,y)$ a vectorului aleator $(\tilde{x},\tilde{y})$ este

$$f(x,y)=f_N(x\mid 0,\sigma^2)\,f_{\chi^2}(y\mid\nu,\sigma^2).$$

Densitatea de probabilitate $f(z,y)$ a vectorului aleator $(\tilde{z},\tilde{y})$ se află pe calea obișnuită, ținînd seama că

$$x=z\sqrt{y/\nu};\quad y=y$$

Scanned with OKEN Scanner

și deci

$$\frac{D(x, y)}{D(z, y)} = \begin{vmatrix} \sqrt{y/\nu} & z/2\sqrt{\nu y} \\ 0 & 1 \end{vmatrix} = \sqrt{y/\nu}.$$

Rezultă atunci că

$$f(z, y) = f_N(z\sqrt{y/\nu} \mid 0, \sigma^2)\, f_{\chi^2}(y \mid \nu, \sigma^2)\sqrt{y/\nu}.$$

Densitatea de probabilitate a lui $\tilde{z}$ este deci

$$f(z) = \int_0^\infty f(z, y)\, dy$$

care, după calcule elementare, se dovedește a fi tocmai $f_{S^*}(z|\nu)$ și teorema este astfel demonstrată.

Din această teoremă și din teorema 13.2.1 rezultă următoarea consecință.

Teorema 13.4.3. *Dacă* $\tilde{x}_0, \tilde{x}_1, \ldots, \tilde{x}_\nu$ *sînt variabile aleatoare independente din clasa* $N(0, \sigma^2)$, *atunci variabila aleatoare*

$$\tilde{y} = \frac{\tilde{x}_0}{\sqrt{(\tilde{x}_1^2 + \ldots + \tilde{x}_\nu^2)/\nu}} \tag{13.4.7}$$

aparține clasei $S^*(\nu)$.

Un alt rezultat interesant în această direcție este următorul.

Teorema 13.4.4. *Fie* $\tilde{x}_1, \ldots, \tilde{x}_\nu$ *variabile aleatoare independente din clasa* $N(0, \sigma^2)$ *și fie*

$$\tilde{y} = (\tilde{x}_1 + \ldots + \tilde{x}_\nu)/\nu \tag{13.4.8}$$

În aceste condiții, variabila aleatoare

$$\tilde{z} = \frac{\tilde{y}\sqrt{\nu(\nu-1)}}{\sqrt{\sum_{i=1}^{\nu} (\tilde{x}_i - \tilde{y})^2}} \tag{13.4.9}$$

aparține clasei $S^*(\nu-1)$.

Demonstrație. Să considerăm variabilele aleatoare $\tilde{y}_1, \ldots, \tilde{y}_\nu$ definite în modul următor:

$$\tilde{y}_i = \sum_{j=1}^{\nu} a_{ij}\tilde{x}_j, \quad 1 \leqslant i \leqslant \nu - 1, \tag{13.4.10}$$

$$\tilde{y}_\nu = \frac{1}{\sqrt{\nu}}(\tilde{x}_1 + \ldots + \tilde{x}_\nu), \tag{13.4.11}$$

unde coeficienții sînt aleși astfel încît matricea pe care o formează să fie ortogonală. Deoarece $\tilde{x}_1, \ldots, \tilde{x}_\nu$ sînt independente din clasa $N(0, \sigma^2)$, rezultă că $\tilde{y}_1, \ldots, \tilde{y}_\nu$ sînt de asemenea independente și aparțin clasei $N(0, \sigma^2)$. Mai mult, rezultă că

$$\tilde{y}_1^2 + \ldots \tilde{y}_\nu^2 = \tilde{x}_1^2 + \ldots + \tilde{x}_\nu^2, \tag{13.4.12}$$

de unde

$$\begin{aligned}\tilde{y}_1^2 + \ldots + \tilde{y}_{\nu-1}^2 &= \tilde{x}_1^2 + \ldots \tilde{x}_\nu^2 - \frac{(\tilde{x}_1 + \ldots + \tilde{x}_\nu)^2}{\nu} = \\ &= \tilde{x}_1^2 + \ldots + \tilde{x}_\nu^2 - \nu\tilde{y}^2 = \sum_{i=\nu}^{n}(\tilde{x}_i - \tilde{y})^2.\end{aligned} \tag{13.4.13}$$

Utilizînd relațiile (13.4.8), (13.4.9), (13.4.11) și (13.4.13), rezultă că putem scrie

$$\tilde{z} = \frac{\frac{1}{\nu}\sqrt{\nu(\nu-1)}\,(\tilde{x}_1 + \ldots + \tilde{x}_\nu)}{\sqrt{\sum_{i=1}^{\nu}(\tilde{x}_i - \tilde{y})^2}} = \frac{\frac{1}{\nu}\sqrt{\nu(\nu-1)}\cdot\sqrt{\nu}\cdot\tilde{y}_\nu}{\sqrt{\sum_{i=1}^{\nu-1}\tilde{y}_i^2}} =$$

$$= \frac{\tilde{y}_\nu}{\sqrt{(\tilde{y}_1^2 + \ldots + \tilde{y}_{\nu-1}^2)/(\nu-1)}}$$

și deci, conform teoremei precedente, rezultă $\tilde{z} \in S^*(\nu-1)$.

Este util să prezentăm și repartiția Student nestandardizată. Vom spune că o variabilă aleatoare are *repartiție Student (nestan-*

dardizată) de parametrii $m \in R$, $\sigma^2 > 0$ *și* $\nu > 0$ dacă densitatea de probabilitate este de forma

$$f_S(x \mid m, \sigma^2, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sigma\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}\left[1 + \frac{(x-m)^2}{\nu\sigma^2}\right]^{-\frac{\nu+1}{2}}, \quad (13.4.14)$$

definită pentru orice $x \in R$. Mulțimea variabilelor aleatoare cu această repartiție va fi notată prin $S(m, \sigma^2, \nu)$.

Este ușor de văzut că $S(0, 1, \nu) = S^*(\nu)$. Mai mult, se poate de asemenea vedea ușor că variabila aleatoare $\tilde{x} \in S(m, \sigma^2, \nu)$ dacă și numai dacă variabila aleatoare

$$\tilde{y} = (\tilde{x} - m)/\sigma \quad (13.4.15)$$

aparține clasei $S^*(\nu)$. Într-adevăr, funcția de repartiție a lui $\tilde{x}$ se exprimă cu ajutorul funcției de repartiție a lui $\tilde{y}$ în modul următor

$$F_S(x \mid m, \sigma^2, \nu) = F_{S^*}\left(\frac{x-m}{\sigma} \mid \nu\right), \quad (13.4.16)$$

iar densitățile de probabilitate corespunzătoare sînt legate de relația

$$f_S(x \mid m, \sigma^2, \nu) = (1/\sigma)\, f_{S^*}\left(\frac{x-m}{\sigma} \mid \nu\right), \quad (13.4.17)$$

de unde urmează ușor afirmația de mai sus.

Dacă $\tilde{x} \in S(m, \sigma^2, \nu)$, atunci, ținînd seama de (13.4.15), rezultă că media sa este

$$M(\tilde{x}) = M(\sigma\tilde{y} + m) = m + \sigma\, M(\tilde{y}) = m, \quad (13.4.18)$$

cu condiția ca $\nu > 1$, iar dispersia

$$V(\tilde{x}) = V(\sigma\tilde{y} + m) = \sigma^2 V(\tilde{y}) = \frac{\sigma^2\nu}{\nu - 2}, \quad (13.4.19)$$

cu condiția ca $\nu > 2$.

Scanned with OKEN Scanner

O altă proprietate interesantă a repartiției Student este dată în teorema următoare, care reprezintă în același timp o posibilitate de definire a acestei repartiții.

Teorema 13.4.3. *Fie* $\tilde{x}$ *o variabilă aleatoare din clasa* $\gamma 2(\psi, \nu)$. *Dacă* $\tilde{y}$ *este o variabilă aleatoare continuă a cărei densitate de probabilitate condiționată de* $\tilde{x} = x$ *este* $f_N(y \mid m, \sigma^2/x)$, *atunci rezultă că* $\tilde{y} \in S(m, \sigma^2 \psi, \nu)$.

Demonstrație. Densitatea de probabilitate $f(x, y)$ a vectorului aleator $(\tilde{x}, \tilde{y})$ se poate scrie sub forma

$$f(x, y) = f_{\gamma_2}(x \mid \psi, \nu) f_N(y \mid m, \sigma^2/x),$$

ținînd seama de ipotezele teoremei. Rezultă că densitatea de probabilitate a lui $\tilde{y}$ este

$$f(y) = \int_0^\infty f_{\gamma_2}(x \mid \psi, \nu) f_N(y \mid m, \sigma^2/x)\, dx =$$

$$= \int_0^\infty \frac{\left(\frac{\psi\nu}{2}\right)^{\frac{\nu}{2}} e^{-\frac{\psi\nu x}{2}} x^{\frac{\nu}{2}-1}}{\Gamma\left(\frac{\nu}{2}\right)} \cdot \frac{\sqrt{x}}{\sigma\sqrt{2\pi}} e^{-\frac{x(y-m)^2}{2\sigma^2}} dx =$$

$$= \frac{\left(\frac{\psi\nu}{2}\right)^{\frac{\nu}{2}}}{\sigma\sqrt{2\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \int_0^\infty e^{-\frac{x}{2}\left[\psi\nu + \left(\frac{y-m}{\sigma}\right)^2\right]} x^{\frac{\nu-1}{2}} dx.$$

Dacă în ultima integrală se face schimbarea de variabila

$$z = \frac{x}{2}\left[\psi\nu + \frac{(y-m)^2}{\sigma^2}\right],$$

atunci densitatea de probabilitate căutată devine

$$\frac{\left(\frac{\psi\nu}{2}\right)^{\frac{\nu}{2}}}{\sigma\sqrt{2\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\int_0^\infty e^{-z}\,z^{\frac{\nu}{2}-1}\left[\frac{2}{\psi\nu+\frac{(y-m)^2}{\sigma^2}}\right]^{\frac{\nu+1}{2}}dz=$$

$$=\frac{(\psi\nu)^{\frac{\nu}{2}}\,\Gamma\left(\frac{\nu+1}{2}\right)}{\sigma\sqrt{\pi}\,\Gamma\left(\frac{\nu}{2}\right)\left[\psi\nu+\left(\frac{y-m}{\sigma}\right)^2\right]^{\frac{\nu+1}{2}}}=$$

$$=\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sigma\sqrt{\psi}\,\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}\left[1+\frac{(y-m)^2}{\psi\nu\sigma^2}\right]^{-\frac{\nu+1}{2}}=f_S(y\mid m,\sigma^2\psi,\nu),$$

adică tocmai faptul că $\tilde{y}\in S(m,\sigma^2\psi,\nu)$.

Să observăm că, fără a restrînge generalitatea teoremei, ne putem mărgini la cazul $\psi=1$.

Din teorema 13.4.3 decurge ușor consecința următoare.

Teorema 13.4.4. *Fie $\tilde{x}$ o variabilă aleatoare din clasa $\gamma 2(\psi,\nu)$. Fie $\tilde{y}_1,\ldots,\tilde{y}_\nu$ variabile aleatoare independente avînd densitățile de probabilitate condiționate de $\tilde{x}=x$ de forma $f_N(y_i\mid m_i,\sigma_i^2/x)$, $1\leqslant i\leqslant\nu$, și fie*

$$\tilde{y}=a_0+\sum_{i=1}^{\nu}a_i\tilde{y}_i. \tag{13.4.20}$$

În aceste condiții, variabila aleatoare $\tilde{y}$ aparține clasei $S(m,\sigma^2\psi,\nu)$, unde

$$m=a_0+\sum_{i=1}^{\nu}a_im_i \tag{13.4.21}$$

și

$$\sigma^2=\sum_{i=1}^{\nu}a_i^2\sigma_i^2. \tag{13.4.22}$$

Demonstrație. Conform unui rezultat cunoscut, variabila aleatoare $\tilde{y}$ dată de (13.4.20) are densitatea de probabilitate condiționată de $\tilde{x} = x$ chiar densitatea de probabilitate normală $f_N(y \mid m, \sigma^2/x)$, unde m și σ^2 sînt dați de formulele (13.4.21) și (13.4.22). Aplicînd teorema precedentă, rezultă că $\tilde{y} \in S(m, \sigma^2 \psi, \nu)$, ceea ce dealtfel trebuia demonstrat.

Repartiția $S(m, \sigma^2, 1)$ mai este numită și *repartiția Cauchy*, pe care o vom nota prin $C(m, \sigma^2)$. Densitatea de probabilitate a unei repartiții Cauchy este deci

$$f_C(x \mid m, \sigma^2) = \frac{\sigma}{\pi} [\sigma^2 + (x - m)^2]^{-1}.$$

Această repartiție nu are momente de nici un ordin. Densitatea de probabilitate $f_C(x \mid m, \sigma^2)$ are graficul în formă de clopot, simetric față de $x = m$, care reprezintă modul și mediana acestei repartiții, iar valoarea maximă este $1/\pi\sigma$.

Funcția caracteristică corespunzătoare este

$$\varphi_C(x \mid m, \sigma^2) = e^{itm - \sigma|t|}.$$

Se vede ușor că, dacă variabilele aleatoare $\tilde{x}_1, \ldots, \tilde{x}_\nu$ sînt independente și $\tilde{x}_j \in C(m_j, \sigma_j^2)$, $1 \leqslant j \leqslant \nu$, atunci variabila aleatoare

$$y = a_0 + \sum_{j=1}^{\nu} a_j \tilde{x}_j$$

aparține de asemenea clasei $C(m, \sigma^2)$, unde

$$m = a_0 + \sum_{j=1}^{\nu} a_j m_j$$

și

$$\sigma = \sum_{j=1}^{\nu} \sigma_j \mid a_j \mid.$$

Într-adevăr,

$$\varphi_{\tilde{\nu}}(t) = \mathrm{M}\left(\mathrm{e}^{\mathrm{i}ta_0}\prod_{j=1}^{\nu}\mathrm{e}^{\mathrm{i}ta_j\tilde{x}_j}\right) = \prod_{j=1}^{\nu}(\mathrm{e}^{\mathrm{i}tm_ja_j-\sigma_j|t|a_j})\mathrm{e}^{\mathrm{i}ta_0} =$$

$$= \mathrm{e}^{\mathrm{i}tm-\sigma|t|} = \varphi_C(t \mid m, \sigma^2),$$

ceea ce demonstrează afirmația.

În particular, dacă $\tilde{x}_1, \ldots, \tilde{x}_\nu$ sînt independente și identic repartizate din clasa $C(m, \sigma^2)$, rezultă că media lor aritmetică

$$(\tilde{x}_1 + \ldots + \tilde{x}_\nu)/\nu$$

aparține de asemenea clasei $C(m, \sigma^2)$.

13.5. REPARTIȚIA BETA MULTIDIMENSIONALĂ

În § 10.3 am prezentat pe larg repartiția normală multidimensională. Vom considera acum alte repartiții multidimensionale continue mai importante în aplicații.

Să considerăm în R^k domeniul D definit de

$$D = \left\{\mathbf{x} = (x_1, \ldots, x_k) : x_i > 0,\ 1 \leqslant i \leqslant k,\ \sum_{i=1}^{k} x_i < 1\right\}. \qquad (13.5.1)$$

Vom spune că vectorul aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_k)'$ are *repartiție beta cu parametrii* $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_k)' > 0$ *și* $\nu > 0$ dacă densitatea de probabilitate a lui $\tilde{\mathbf{x}}$ este

$$f_\beta(\mathbf{x} \mid \boldsymbol{\rho}, \nu) = \begin{cases} \dfrac{x_1^{\rho_1-1} \ldots x_k^{\rho_k-1}(1-x_1-\ldots-x_k)^{\nu-1}}{B_k(\rho, \nu)} & \text{pentru } \mathbf{x} \in D, \\ 0 & \text{pentru } \mathbf{x} \notin D, \end{cases} \qquad (13.5.2)$$

unde

$$B_k(\rho, \nu) = \int_D x_1^{\rho_1 - 1} \dots x_k^{\rho_k - 1} (1 - x_1 - \dots - x_k)^{\nu - 1} \, dx_1 \dots dx_k \quad (13.5.3)$$

este *funcția beta k—dimensională sau funcția lui Dirichlet;* o prezentare succintă a proprietăților acestei funcții este dată în Anexă (vezi § *A*.3).

Mulțimea vectorilor aleatori cu această repartiție va fi notată prin $\beta(\rho, \nu)$. Evident, pentru $k = 1$ se obține repartiția beta unidimensională.

Se poate de asemenea arăta ușor (vezi § *A*.3) că funcția lui Dirichlet $B_k(\rho, \nu)$ se poate exprima cu ajutorul funcției gama în modul următor

$$B_k(\rho, \nu) = \frac{\Gamma(\rho_1)\,\Gamma(\rho_2) \dots \Gamma(\rho_k)\,\Gamma(\nu)}{\Gamma(\rho_1 + \rho_2 + \dots + \rho_k + \nu)}. \quad (13.5.4)$$

După cum am văzut, repartiția beta unidimensională este utilizată ca instrument de exprimare a unor opinii privind o probabilitate necunoscută (de exemplu, proporția pieselor defecte dintr-un lot de produse). Repartiția beta k—dimensională este utilizată în situații analoage, dar mai complicate. Mai precis, ea este utilizată pentru a exprima opinii asupra unor proporții necunoscute $p_1, \dots, p_k$ relative la o populație care este formată din $k+1$ clase disjuncte $C_0, C_1, \dots, C_k$, unde C_0 este clasa care conține proporția $p_0 = 1 - (p_1 + \dots + p_k)$ din populație.

Momentele repartiției beta $\beta(\rho, \nu)$ se calculează ușor dacă observăm că momentul de ordinul $\mathbf{r} = (r_1, \dots, r_k)'$ al vectorului aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_k)' \in \beta(\rho, \nu)$ este

$$M_{r_1 \dots r_k}(\tilde{\mathbf{x}}) =$$

$$= \frac{1}{B_k(\rho, \nu)} \int_D x_1^{r_1 + \rho_1 - 1} x_2^{r_2 + \rho_2 - 1} \dots x_k^{r_k + \rho_k - 1} (1 - x_1 - \dots - x_k)^{\nu - 1} dx_1 \dots dx_k$$

(13.5.5.)

și, ținînd seama de (13.5.3), rezultă

$$M_{r_1 \dots r_k}(\tilde{\mathbf{x}}) = \frac{B_k(\boldsymbol{\rho} + \mathbf{r}, \nu)}{B_k(\boldsymbol{\rho}, \nu)}. \tag{13.5.6}$$

Utilizînd (13.5.4), rezultă că (13.5.6) se scrie sub forma

$$M_{r_1 \dots r_k}(\tilde{\mathbf{x}}) = \frac{\Gamma(\rho_1+r_1) \dots \Gamma(\rho_k+r_k)\, \Gamma(\rho_1+\dots+\rho_k+\nu)}{\Gamma(\rho_1)\dots\Gamma(\rho_k)\, \Gamma(\rho_1+\dots+\rho_k+r_1+\dots+r_k+\nu)}. \tag{13.5.7}$$

Atunci

$$M(\tilde{x}_j) = \frac{\Gamma(\rho_j+1)\, \Gamma(\rho_1+\dots+\rho_k+\nu)}{\Gamma(\rho_j)\, \Gamma(\rho_1+\dots+\rho_k+\nu+1)} = \frac{\rho_j}{\rho_1+\dots+\rho_k+\nu}, \tag{13.5.8}$$

$$M_2(\tilde{x}_j) = \frac{\rho_j(\rho_j+1)}{(\rho_1+\dots+\rho_k+\nu)(\rho_1+\dots+\rho_k+\nu+1)}, \tag{13.5.9}$$

$$V(\tilde{x}_j) = \frac{\rho_j(\rho_1 + \dots + \rho_{j-1} + \rho_{j+1} + \dots + \rho_k + \nu)}{(\rho_1 + \dots + \rho_k+\nu)^2(\rho_1 + \dots + \rho_k + \nu + 1)}, \tag{13.5.10}$$

$$M(\tilde{x}_i \tilde{x}_j) = \frac{\rho_i \rho_j}{(\rho_1+\dots+\rho_k+\nu)(\rho_1+\dots+\rho_k+\nu+1)}, \quad i \neq j \tag{13.5.11}$$

și

$$\operatorname{cov}(\tilde{x}_i, \tilde{x}_j) = \frac{-\rho_i \rho_j}{(\rho_1 + \dots + \rho_k+\nu)^2(\rho_1+ \dots + \rho_k+\nu+1)}. \tag{13.5.12}$$

Să considerăm un vector aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_k)'$ din clasa $\beta(\boldsymbol{\rho}, \nu)$ și fie $(\tilde{\mathbf{x}}_1', \tilde{\mathbf{x}}_2')'$ o partiție a sa, unde $\tilde{\mathbf{x}}_1 = (\tilde{x}_{i_1} \dots, \tilde{x}_{i_{k_1}})'$ cu $0 < k_1 < k$. Dacă partiționăm vectorul $\boldsymbol{\rho}$ în mod corespunzător în $(\boldsymbol{\rho}_1', \boldsymbol{\rho}_2')'$ atunci ne punem mai întîi problema determinării repartiției (marginale) a lui $\tilde{\mathbf{x}}_1$ și a lui $\tilde{\mathbf{x}}_2$. În acest sens avem rezultatul care urmează.

Teorema 13.5.1. *Dacă* $\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_k)' \in \beta(\boldsymbol{\rho}, \nu)$, *atunci*

$$\tilde{\mathbf{x}}_1 = (\tilde{x}_{i_1}, \dots, \tilde{x}_{i_{k_1}})' \in \beta(\boldsymbol{\rho}_1, \nu^*), \text{ unde } \boldsymbol{\rho}_1 = (\rho_{i_1}, \dots, \rho_{i_{k_1}})'$$

și

$$\nu^* = \sum_{i \notin \{i_1, \dots, i_{k_1}\}} \rho_i + \nu. \tag{13.5.13}$$

Demonstrație. Putem presupune fără a restrînge generalitatea că $\tilde{\mathbf{x}}_1 = (\tilde{x}_1, \ldots, \tilde{x}_{k_1})'$, renumerotînd eventual componentele lui $\tilde{\mathbf{x}}$; pentru simplitatea scrierii vom nota $k_1 = r$.

Pentru $(x_1, \ldots, x_r)$ fixat definim domeniul

$$D_0 = \left\{ (x_{r+1}, \ldots, x_k) : \sum_{i=r+1}^{k} x_i < x^*, \; x_i > 0, \; r+1 \leqslant i \leqslant k \right\},$$

unde

$$x^* = 1 - \sum_{i=1}^{r} x_i.$$

Densitatea de probabilitate $f(x_1, \ldots, x_r)$ a vectorului aleator $\tilde{\mathbf{x}}_1 = (\tilde{x}_1, \ldots, \tilde{x}_r)'$ se poate exprima în modul următor

$$f(x_1, \ldots, x_r) = \int_{D_0} f_\beta(\mathbf{x} \mid \boldsymbol{\rho}, \nu) \, \mathrm{d}x_{r+1} \ldots \mathrm{d}x_k =$$

$$= \frac{1}{B_k(\boldsymbol{\rho}, \nu)} \int_{D_0} x_1^{\rho_1 - 1} \ldots x_k^{\rho_k - 1} (1 - x_1 - \ldots - x_k)^{\nu - 1} \, \mathrm{d}x_{r+1} \ldots \mathrm{d}x_k =$$

$$= \frac{x_1^{\rho_1 - 1} \ldots x_1^{\rho_k - 1}}{B_k(\boldsymbol{\rho}, \nu)} \int_{D_0} x_{r+1}^{\rho_{r+1} - 1} \ldots x_k^{\rho_k - 1} \left(x^* - \sum_{i=r+1}^{k} x_i \right)^{\nu - 1} \mathrm{d}x_{r+1} \ldots \mathrm{d}x_k.$$

Făcînd schimbarea de variabile

$$x_i = y_i \cdot x^*, \quad r+1 \leqslant i \leqslant k,$$

obținem

$$f(x_1, \ldots, x_r) =$$

$$= \frac{x_1^{\rho_1 - 1} \ldots x_r^{\rho_k - 1} (x^*)^{\nu^* - 1}}{B_k(\boldsymbol{\rho}, \nu)} \int_{D^*} y_{r+1}^{\rho_{r+1} - 1} \ldots$$

$$\ldots y_k^{\rho_k - 1} \left(1 - \sum_{i=r+1}^{k} y_i \right)^{\nu - 1} \mathrm{d}y_{r+1} \ldots \mathrm{d}y_k,$$

unde ν^* este dat în enunțul teoremei iar

$$D^*=\left\{(y_{r+1},\ldots,y_k)':\sum_{i=r+1}^{k}y_i<1,\ y_i>0,\ r+1\leqslant i\leqslant k\right\}.$$

Rezultă atunci că

$$f(x_1,\ldots,x_r)=\frac{x_1^{\rho_1-1}\ldots x_r^{\rho_r-1}(1-x_1-\ldots-x_r)^{\nu^*-1}}{B_k(\boldsymbol{\rho},\ \nu)}B_{k-r}(\boldsymbol{\rho}_2,\ \nu). \tag{13.5.14}$$

În sfîrșit, conform relației (13.5.4), putem scrie

$$\frac{B_{k-r}(\boldsymbol{\rho}_2,\ \nu)}{B_k(\boldsymbol{\rho},\ \nu)}=\frac{\Gamma(\rho_{r+1})\ldots\Gamma(\rho_k)\,\Gamma(\nu)}{\Gamma(\rho_{r+1}+\ldots+\rho_k+\nu)}\cdot\frac{\Gamma(\rho_1+\ldots+\rho_k+\nu)}{\Gamma(\rho_1)\ldots\Gamma(\rho_k)\Gamma(\nu)}=$$

$$=\frac{\Gamma(\rho_1+\ldots+\rho_k+\nu^*)}{\Gamma(\rho_1)\ldots\Gamma(\rho_k)\,\Gamma(\nu^*)}=\frac{1}{B_k(\boldsymbol{\rho}_1,\ \nu^*)}. \tag{13.5.15}$$

Din (13.5.14) și (13.5.15) rezultă că densitatea de probabilitate $f(x_1,\ldots,x_r)$ coincide cu $f_\beta(x_1,\ldots,x_r\mid\boldsymbol{\rho}_1,\nu^*)$ și astfel teorema este demonstrată.

Din această teoremă obținem imediat densitatea de probabilitate condiționată $f(\mathbf{x}_2\mid\mathbf{x}_1)$. Într-adevăr,

$$f(\mathbf{x}_2\mid\mathbf{x}_1)=\frac{f_\beta(\mathbf{x}\mid\boldsymbol{\rho},\nu)}{f_\beta(\mathbf{x}_1\mid\boldsymbol{\rho}_1,\nu^*)}$$

și deci, utilizînd rezultatul precedent, obținem

$$f(\mathbf{x}_2\mid\mathbf{x}_1)=\frac{B_r(\boldsymbol{\rho}_1,\ \nu^*)}{B_k(\boldsymbol{\rho},\ \nu)}\cdot\frac{(x_{r+1})^{\rho_{r+1}-1}\ldots x_k^{\rho_k-1}(1-x_1-\ldots-x_k)^{\nu-1}}{(1-x_1-\ldots-x_k)^{\nu^*-1}}. \tag{13.5.16}$$

Dacă notăm

$$x^*(r)=1-\sum_{i=1}^{r}x_i \tag{13.5.17}$$

și considerăm vectorul aleator

$$\tilde{\mathbf{x}}_2^* = \frac{1}{x^*(r)}\tilde{\mathbf{x}}_2, \tag{13.5.18}$$

rezultă imediat că densitatea de probabilitate $f^*(\mathbf{x}_2 \mid \mathbf{x}_1)$ a lui $\tilde{\mathbf{x}}_2^*$ condiționată de $\tilde{\mathbf{x}}_1 = \mathbf{x}_1$ este

$$f^*(\mathbf{x}_2 \mid \mathbf{x}_1) = \frac{B_r(\boldsymbol{\rho}_1, \nu^*)}{B_k(\boldsymbol{\rho}, \nu)} x_{r+1}^{\rho_{r+1}-1} \dots x_k^{\rho_k-1}(1 - x_1 - \dots - x_k)^{\nu}$$

sau, ținînd seama de (13.5.15), rezultă că

$$f^*(\mathbf{x}_2 \mid \mathbf{x}_1) = \frac{x_{r+1}^{\rho_{r+1}} \dots x_k^{\rho_k}(1 - x_1 - \dots - x_k)^{\nu}}{B_{k-r}(\boldsymbol{\rho}_2, \nu)}, \tag{13.5.19}$$

ceea ce arată că $f^*(\mathbf{x}_2 \mid \mathbf{x}_1)$ coincide cu $f_\beta(\mathbf{x}_2 \mid \boldsymbol{\rho}_2, \nu)$.

Am obținut deci rezultatul următor.

Teorema 13.5.2. *Dacă* $\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_k)' \in \beta(\boldsymbol{\rho}, \nu)$ *și* $(\tilde{\mathbf{x}}_1', \tilde{\mathbf{x}}_2')'$ *este o partiție a lui* $\tilde{\mathbf{x}}$, $(\boldsymbol{\rho}_1', \boldsymbol{\rho}_2')'$ *este partiția corespunzătoare a lui* $\boldsymbol{\rho}$, *iar* $\tilde{\mathbf{x}}_2^*$ *este vectorul aleator definit de* (13.5.18), *atunci densitatea de probabilitate a lui* $\tilde{\mathbf{x}}_2^*$ *condiționată de* $\tilde{\mathbf{x}}_1 = \mathbf{x}_1$ *este* $f_\beta(\cdot \mid \boldsymbol{\rho}_2, \nu)$.

Este clar că densitatea de probabilitate $f(\mathbf{x}_2 \mid \mathbf{x}_1)$ a vectorului aleator $\tilde{\mathbf{x}}_2$ condiționată de $\tilde{\mathbf{x}}_1 = \mathbf{x}_1$, care, conform relațiilor (13.5.15), (13.5.16) și (13.5.17), are forma

$$f(\mathbf{x}_2 \mid \mathbf{x}_1) = \frac{x_{r+1}^{\rho_{r+1}-1} \dots x_k^{\rho_k-1}(x^*(r) - x_{r+1} - \dots - x_k)^{\nu-1}}{B_{k-r}(\boldsymbol{\rho}_2, \nu)(x^*(r))^{\nu^*-1}} \tag{13.5.20}$$

nu este densitatea de repartiție beta, dar ea poate fi considerată ca o densitate de probabilitate a repartiției beta *nestandardizată*, asupra căreia nu insistăm aici.

Vom încheia considerațiile privind repartiția beta multidimensională cu rezultatul care urmează.

Scanned with OKEN Scanner

Teorema 13.5.3. *Dacă* $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_k) \in \beta(\boldsymbol{\rho}, \nu)$,

$$\tilde{y} = \tilde{x}_1 + \ldots + \tilde{x}_k \tag{13.5.21}$$

și

$$\rho = \rho_1 + \ldots + \rho_k, \tag{13.5.22}$$

atunci $\tilde{y} \in \beta(\rho, \nu)$.

Demonstrație. Să considerăm variabila aleatoare

$$\tilde{x}_0 = 1 - (\tilde{x}_1 + \ldots + \tilde{x}_k)$$

și să determinăm momentul de ordinul r al acestei variabile aleatoare:

$$M(\tilde{x}_0^r) = \frac{1}{B_k(\boldsymbol{\rho}, \nu)} \int_D x_1^{\rho_1 - 1} \ldots x_k^{\rho_k - 1} (1 - x_1 - \ldots - x_k)^{r+\nu-1} \, dx_1 \ldots dx_k,$$

unde D este domeniul definit de (13.5.1). Rezultă atunci că avem

$$M(\tilde{x}_0^r) = \frac{B_k(\boldsymbol{\rho}, r+\nu)}{B_k(\boldsymbol{\rho}, \nu)}$$

și, utilizînd din nou (13.5.4), obținem

$$M(\tilde{x}_0^r) = \frac{\Gamma(\rho+\nu)\,\Gamma(r+\nu)}{\Gamma(\nu)\,\Gamma(\nu+\rho+r)} = \frac{B(\nu+r, \rho)}{B(\nu, \rho)},$$

ceea ce arată că momentele de ordinul r ale lui $\tilde{x}_0$ sînt tocmai momentele de ordinul r ale repartiției $\beta(\nu, \rho)$ și deci $\tilde{x}_0 \in \beta(\nu, \rho)$, ceea ce este echivalent cu $\tilde{y} = 1 - \tilde{x}_0 \in \beta(\rho, \nu)$, și astfel teorema este demonstrată.

Scanned with OKEN Scanner

13.6. REPARTIȚIA STUDENT MULTIDIMENSIONALĂ

Spunem că vectorul aleator $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_k)'$ are *repartiție Student (nesingulară) de parametrii* $\mathbf{m} \in R^k$, $\boldsymbol{\delta}^2$ *(matrice simetrică pozitiv definită) și* $\nu > 0$ dacă densitatea sa de probabilitate este

$$f_S(\mathbf{x} \mid \mathbf{m}, \boldsymbol{\delta}^2, \nu) =$$

$$= \frac{\Gamma\left(\frac{\nu + k}{2}\right)}{(\det \boldsymbol{\delta}^2)^{\frac{1}{2}} (\pi\nu)^{\frac{k}{2}} \Gamma\left(\frac{\nu}{2}\right)} \left[1 + \frac{1}{\nu} (\mathbf{x} - \mathbf{m})' \boldsymbol{\delta}^{-2} (\mathbf{x} - \mathbf{m})\right],$$

fiind definită pentru orice $\mathbf{x} \in R^k$, unde $\boldsymbol{\delta}^{-2} = (\boldsymbol{\delta}^2)^{-1}$. Mulțimea vectorilor aleatori cu această repartiție va fi notată prin $S(\mathbf{m}, \delta^2, \nu)$.

Se poate da un rezultat analog cu teorema 13.4.3, valabilă în cazul unidimensional.

Teorema 13.6.1. *Dacă* $\tilde{x} \in \gamma 2(\psi, \nu)$ *și dacă* $\tilde{\mathbf{y}} = (\tilde{y}_1, \ldots, \tilde{y}_k)'$ *este un vector aleator a cărui densitate de probabilitate condiționată de* $\tilde{\mathbf{x}} = \mathbf{x}$ *este* $\mathrm{f_N}\left(\mathbf{y} \mid \mathbf{m}, \frac{1}{\mathrm{x}} \boldsymbol{\delta}^2\right)$, *unde* $\mathbf{m} \in \mathrm{R^k}$ *și* δ^2 *este matrice simetrică pozitiv definită de ordinul* k, *atunci avem* $\tilde{\mathbf{y}} \in \mathrm{S}(\mathbf{m}, \psi \delta^2, \nu)$.

Demonstrație. Să considerăm vectorul aleator $(\tilde{x}, \tilde{y}_1, \ldots, \tilde{y}_k)'$. Densitatea de probabilitate a acestui vector este

$$f(x, y_1, \ldots, y_k) = f_{\gamma 2}(x \mid \psi, \nu) f_N\left(y_1, \ldots, y_k \mid \mathbf{m}, \frac{1}{x} \boldsymbol{\delta}^2, \nu\right). \tag{13.6.2}$$

Rezultă atunci că densitatea de probabilitate $f(y_1, \ldots, y_k)$ a vectorului aleator $\tilde{\mathbf{y}}$ este dată de relația

$$f(y_1, \ldots, y_k) = \int_0^\infty f_{\gamma 2}(x \mid \psi, \nu) f_N\left(y_1, \ldots, y_k \mid \mathbf{m}, \frac{1}{x} \boldsymbol{\delta}^2, \nu\right) dx,$$

Scanned with OKEN Scanner

adică

$$f(y_1, \ldots, y_k) =$$

$$= \int_0^\infty \frac{\left(\frac{\psi\nu}{2}\right)^{\frac{\nu}{2}} e^{-\frac{\psi\nu x}{2}} x^{\frac{\nu}{2}-1}}{\Gamma\left(\frac{\nu}{2}\right)} \cdot \frac{1}{(2\pi)^{\frac{k}{2}} \sqrt{\det\left(\frac{1}{x}\,\delta^2\right)}} \exp\left[-\frac{1}{2}(y - \right.$$

$$\left. - m)' \delta^{-2}(y - m)\right] dx =$$

$$= \frac{\left(\frac{\psi\nu}{2}\right)^{\frac{\nu}{2}}}{(2\pi)^{\frac{k}{2}} \Gamma\left(\frac{\nu}{2}\right) \sqrt{\det(\delta^2)}} \int_0^\infty e^{-\frac{x}{2}[\psi\nu + (y-m)'\delta^{-2}(y-m)]} x^{\frac{k+\nu}{2}-1} dx.$$

Dacă în ultima integrală se face schimbarea de variabilă

$$z = \frac{x}{2}(\psi\nu + (y - m)' \delta^{-2}(y - m)),$$

atunci

$$f(y_1, \ldots, y_k) =$$

$$= \frac{\left(\frac{\psi\nu}{2}\right)^{\frac{\nu}{2}} \Gamma\left(\frac{\nu + k}{2}\right)}{\pi^{\frac{k}{2}} \Gamma\left(\frac{\nu}{2}\right) \sqrt{\det \delta^2}} [\psi\nu + (y - m)' \delta^{-2}(y - m)]^{-\frac{k+\nu}{2}} =$$

$$= \frac{\Gamma\left(\frac{\nu + k}{2}\right)}{(\pi\nu)^{\frac{k}{2}} \Gamma\left(\frac{\nu}{2}\right) \sqrt{\det \delta^2}} \left[1 + \frac{1}{\nu}(y-m)'\left(\frac{1}{\psi}\,\delta^{-2}\right)(y - m)\right]^{-\frac{k+\nu}{2}} =$$

$$= f_S(y \mid m, \psi\delta^2, \nu),$$

adică tocmai faptul că $\tilde{y} \in S(m, \psi\delta^2, \nu)$.

Dacă $g : R^k \to R$ este o funcție măsurabilă cu proprietatea că $g(\tilde{y}_1, ..., \tilde{y}_k)$ este integrabilă, rezultă că

$$M(g(\tilde{y}_1, ..., \tilde{y}_k)) = M(M(g(\tilde{y}_1, ..., \tilde{y}_k) \mid \tilde{x}), \qquad (13.6.3)$$

ceea ce ne permite să transcriem cu ușurință rezultatele obținute pentru repartiția normală multidimensională în cazul repartiției Student.

Observație. Să presupunem că vectorul aleator $(\tilde{y}_1, ..., \tilde{y}_k)'$ are repartiția condiționată de variabila aleatoare $\tilde{x}$ dată de repartiția normală singulară de parametrii $\mathbf{m}$ și $\frac{1}{\tilde{x}}\ \boldsymbol{\delta}^2$. Se poate atunci defini repartiția Student singulară de parametrii $\mathbf{m}$, $\boldsymbol{\delta}^2$ și $\nu > 0$, ca fiind repartiția (necondiționată) a vectorului aleator $(\tilde{y}_1, ..., \tilde{y}_k)'$.

Să aplicăm relația (13.6.3) stabilită mai sus pentru calculul mediei și matricei de covarianță a unui vector aleator $\tilde{\mathbf{y}} = (\tilde{y}_1, ..., \tilde{y}_k)' \in S(\mathbf{m}, \boldsymbol{\delta}^2, \nu)$. Pentru simplitatea exprimării, valoarea medie în raport cu o anumită repartiție se va nota analog cu notația utilizată pînă acum pentru densitatea de probabilitate, funcția de repartiție sau funcția caracteristică. Cu această convenție, rezultă că relația (13.6.3) se poate scrie sub forma

$$M_S(g(\tilde{\mathbf{y}}) \mid \mathbf{m}, \boldsymbol{\delta}^2, \nu) = M_{\gamma 2}\left(M_N\left(g(\tilde{\mathbf{y}}) \mid \mathbf{m}, \frac{1}{\tilde{x}}\ \boldsymbol{\delta}^2\right) \mid 1, \nu\right). \qquad (13.6.4)$$

Rezultă atunci că media lui $\tilde{\mathbf{y}}$ se poate scrie sub forma

$$M(\tilde{\mathbf{y}}) = M_S(\tilde{\mathbf{y}} \mid \mathbf{m}, \boldsymbol{\delta}^2, \nu) = M_{\gamma 2}\left(M_N\left(\tilde{\mathbf{y}} \mid \mathbf{m}, \frac{1}{\tilde{x}}\ \boldsymbol{\delta}^2\right) \mid 1, \nu\right).$$

Deoarece

$$M_N\left(\tilde{\mathbf{y}} \mid \mathbf{m}, \frac{1}{\tilde{x}}\ \boldsymbol{\delta}^2\right) = \mathbf{m},$$

obținem

$$M_{\gamma 2}(\mathbf{m} \mid 1, \nu) = \mathbf{m},$$

de unde

$$M_S(\tilde{\mathbf{y}} \mid \mathbf{m},\ \delta^2, \nu) = \mathbf{m}. \tag{13.6.5}$$

Pentru obținerea matricei de covarianță a lui $\tilde{\mathbf{y}}$, să observăm mai întîi că

$$M_N\left((\tilde{\mathbf{y}} - \mathbf{m})(\tilde{\mathbf{y}} - \mathbf{m})' \mid \mathbf{m},\ \frac{1}{\tilde{x}}\,\delta^2\right) = \frac{1}{\tilde{x}}\,\delta^2$$

și deci

$$M_{\gamma 2}\left(\frac{1}{\tilde{x}}\,\delta^2 \mid 1,\ \nu\right) = M_{\gamma 2}\left(\frac{1}{\tilde{x}} \mid 1,\ \nu\right)\delta^2.$$

Deoarece $\tilde{x} \in \gamma 2(1,\ \nu)$ dacă și numai dacă $\dfrac{1}{\tilde{x}} \in \gamma i\,(1,\ \nu)$, rezultă că putem scrie

$$M_{\gamma 2}(1/\tilde{x} \mid 1, \nu) = M_{\gamma i}(\tilde{x} \mid 1, \nu).$$

În sfîrșit, avem

$$M_{\gamma i}(\tilde{x} \mid 1, \nu) = \frac{\left(\dfrac{\nu}{2}\right)^{\frac{\nu}{2}}}{\Gamma\left(\dfrac{\nu}{2}\right)} \int_0^\infty e^{-\nu/2x}\, x^{\nu/2+1}\, dx,$$

care, după schimbarea de variabilă $\nu/2x = y$, devine

$$M_{\gamma i}(\tilde{x} \mid 1, \nu) = \frac{\dfrac{\nu^2}{4}}{\dfrac{\nu}{2}\,\Gamma\left(\dfrac{\nu}{2}\right)} \int_0^\infty e^{-y}\, y^{\frac{\nu}{2}-2}\, dy =$$

$$= \frac{\dfrac{\nu}{2}\,\Gamma\left(\dfrac{\nu}{2} - 1\right)}{\Gamma\left(\dfrac{\nu}{2}\right)} = \frac{\dfrac{\nu}{2}\,\Gamma\left(\dfrac{\nu}{2} - 1\right)}{\left(\dfrac{\nu}{2} - 1\right)\Gamma\left(\dfrac{\nu}{2} - 1\right)} = \frac{\nu}{\nu - 2},$$

cu condiția ca $\nu > 2$. Obținem în definitiv

$$V(\tilde{y}_1, \ldots, \tilde{y}_k) = \frac{\nu}{\nu - 2} \delta^2. \tag{13.6.6}$$

Putem de asemenea demonstra teoreme analoage cu teoremele 10.3.2, 10.3.3, 10.3.5 și 10.3.6, care se obțin imediat din acestea din urmă prin intermediul relației (13.6.4).

13.7. REPARTIȚIA WISHART

Spunem că o matrice aleatoare $\tilde{\mathbf{X}} = (\tilde{x}_{ij})_{1 \leqslant i, j \leqslant k}$ are *repartiție Wishart de ordinul k cu parametrii* $\boldsymbol{\psi}$ *și* ν, unde $\boldsymbol{\psi} = (\psi_{ij})_{1 \leqslant i, j \leqslant k}$, iar $\nu > 0$, dacă densitatea de probabilitate a lui $\tilde{\mathbf{X}}$ este de forma:

$$f_W(\mathbf{X} \mid \boldsymbol{\psi}, \nu) =$$

$$= \begin{cases} \dfrac{(\det \boldsymbol{\psi})^{\frac{\nu}{2}} (\det \mathbf{X})^{\frac{\nu-k-1}{2}} \exp\left(-\frac{1}{2} \operatorname{ur}(\boldsymbol{\psi} \mathbf{X})\right)}{\pi^{\frac{k(k-1)}{4}} 2^{\frac{k\nu}{2}} \Gamma\left(\frac{\nu}{2}\right) \ldots \Gamma\left(\frac{\nu - k + 1}{2}\right)} & \text{dacă } \mathbf{X} \text{ este pozitiv definită,} \\ 0 & \text{în caz contrar,} \end{cases} \tag{13.7.1}$$

unde $\operatorname{ur}(\boldsymbol{\psi}\mathbf{X})$ este urma matricei $\boldsymbol{\psi}\mathbf{X}$; pentru o prezentare sumară a noțiunii de urmă a unei matrice vezi Anexa, § A.4.

Mulțimea matricelor aleatoare simetrice $\tilde{\mathbf{X}}$ care au această repartiție va fi notată prin $W(\boldsymbol{\psi}, \nu)$.

Dacă observăm că $f_W(\mathbf{X} \mid \boldsymbol{\psi}, \nu)$ are valori nenule numai pentru matrice aleatoare $\tilde{\mathbf{X}}$ care sînt simetrice, adică matrice avînd numai $(k(k+1))/2$ elemente funcțional independente, rezultă că putem privi repartiția Wishart ca fiind repartiția vectorului aleator $\tilde{\mathbf{x}} = (\tilde{x}_{11}, \ldots, \tilde{x}_{1k}, \tilde{x}_{22}, \ldots, \tilde{x}_{2k}, \ldots, \tilde{x}_{kk})'$ cu $k(k+1)/2$ componente.

Scanned with OKEN Scanner

Este clar că pentru $k = 1$ repartiția Wishart se reduce la repartiția gama-2 sau la repartiția hi-pătrat și avem

$$W(\psi, \nu) = \gamma 2(\psi/\nu, \nu) = \chi^2(\nu, 1/\psi).$$

Este de asemenea ușor de arătat că, dacă $\widetilde{\mathbf{X}} \in W(\psi, \nu)$ și $a > 0$, atunci $a\widetilde{\mathbf{X}} \in W\left(\frac{1}{a}\psi, \nu\right)$.

Teorema următoare arată modul în care apare această repartiție.

Teorema 13.7.1. *Dacă* $\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_\nu$ *sînt vectori aleatori independenți din clasa* $\mathrm{N}(\mathbf{0}, \delta^2)$, *atunci matricea aleatoare*

$$\widetilde{\mathbf{Y}} = \sum_{i=1}^{\nu} \widetilde{\mathbf{x}}_i \widetilde{\mathbf{x}}_i' \tag{13.7.2}$$

aparține clasei $\mathrm{W}(\delta^{-2}, \nu)$, *unde* δ^{-2} *este inversa matricei* δ^2, *adică* $\delta^{-2} = (\delta^2)^{-1}$.

Demonstrație. Densitatea de probabilitate a vectorului aleator $(\widetilde{\mathbf{x}}_1', \ldots, \widetilde{\mathbf{x}}_\nu')'$ este

$$f(\mathbf{x}_1, \ldots, \mathbf{x}_\nu) = \prod_{i=1}^{\nu} f_N(\mathbf{x}_i \mid \mathbf{0}, \delta^2) =$$

$$= \prod_{i=1}^{\nu} \frac{1}{(2\pi)^{\frac{k}{2}} \sqrt{\det \delta^2}} \exp\left(-\frac{1}{2}\, \mathbf{x}_i' \delta^{-2} \mathbf{x}_i\right) = \tag{13.7.3}$$

$$= \frac{\exp\left(-\frac{1}{2}\sum_{i=1}^{\nu} \mathbf{x}_i' \delta^{-2} \mathbf{x}_i\right)}{(2\pi)^{k\frac{\nu}{2}} (\det \delta^2)^{\frac{\nu}{2}}} = \frac{\exp\left(-\frac{1}{2}\sum_{j,r=1}^{k} \delta_{jr}^{-2} y_{jr}\right)}{(2\pi)^{k\frac{\nu}{2}} (\det \delta^2)^{\frac{\nu}{2}}},$$

unde δ_{jr}^{-2}, $1 \leqslant j, r \leqslant k$ sînt elementele matricei δ^{-2}, iar y_{jr}, $1 \leqslant j, r \leqslant k$ sînt elementele matricei $\mathbf{Y}$, adică

$$y_{jr} = \sum_{i=1}^{\nu} x_{ij} x_{ir} \tag{13.7.4}$$

și x_{ij}, $1 \leqslant j \leqslant k$, sînt componentele vectorului $\mathbf{x}_i$.

Scanned with OKEN Scanner

Funcția caracteristică a matricei aleatoare $\tilde{\mathbf{Y}}$ este

$$\varphi(\mathbf{t}) = M\left(e^{i\sum_{j,r=1}^{k}\tilde{y}_{jr}\,t_{jr}(1+\delta_{jr})/2}\right), \tag{13.7.5}$$

unde $\tilde{y}_{jr}$ și t_{jr} sînt respectiv componentele matricelor $\tilde{\mathbf{Y}}$ și $\mathbf{t}$, iar δ_{jr} sînt simbolurile lui Krönecker.

Din relațiile (13.7.3), (13.7.4) și (13.7.5) rezultă atunci că putem scrie

$$\varphi(t) = \frac{1}{(2\pi)^{k\frac{\nu}{2}}(\det \boldsymbol{\delta}^2)^{\frac{\nu}{2}}}\int \exp\left(-\frac{1}{2}\left(\sum_{j,r=1}^{k}\delta_{jr}^{-2}y_{jr} - \right.\right.$$

$$\left.\left. - i\sum_{j,r=1}^{k}(1+\delta_{jr})t_{jr}y_{jr}\right)\right)\prod dy_{jr} =$$

$$= \frac{1}{(2\pi)^{k\frac{\nu}{2}}(\det \boldsymbol{\delta}^2)^{\frac{\nu}{2}}}\int \exp\left(-\frac{1}{2}\,\mathrm{ur}(\mathbf{AY})\right)\prod dy_{jr}, \tag{13.7.6}$$

unde $\mathbf{A} = (a_{jr})_{1\leqslant j,\, r\leqslant k}$ este matricea cu elementele

$$a_{jr} = \delta_{jr}^{-2} - i(1+\delta_{jr})t_{jr},\quad 1\leqslant j,\quad r\leqslant k. \tag{13.7.7}$$

După efectuarea integralei din (13.7.6), rezultă că

$$\varphi(t) = \left((\det \boldsymbol{\delta}^2)(\det \mathbf{A})\right)^{-\frac{\nu}{2}}. \tag{13.7.8}$$

Dacă observăm însă că pentru orice m întreg avem

$$\int (\det \mathbf{Y})^m \exp\left(-\frac{1}{2}\,\mathrm{ur}(\boldsymbol{\delta}^{-2}\mathbf{Y})\right)\prod dy_{jr} =$$

$$= \frac{\pi^{\frac{k(k-1)}{4}}\,2^{k\left(m+\frac{k+1}{2}\right)}\,\Gamma\left(m+\frac{k+1}{2}\right)\ldots\Gamma(m+1)}{(\det \boldsymbol{\delta}^{-2})^{m+\frac{k+1}{2}}},$$

rezultă că putem defini în $R^{\frac{k(k+1)}{2}}$ o repartiție prin densitatea de probabilitate avînd valori nenule de forma

$$\frac{(\det \delta^{-2})^{m+\frac{k+1}{2}} \det(\mathbf{Y})^m \exp\left(-\frac{1}{2}\operatorname{ur}(\delta^{-2}\mathbf{Y})\right)}{\pi^{\frac{k(k-1)}{4}} 2^{k\left(m+\frac{k+1}{2}\right)} \Gamma\left(m+\frac{k+1}{2}\right)\ldots\Gamma(m+1)} \tag{13.7.9}$$

pentru $\mathbf{Y} \geqslant 0$.

Se poate vedea ușor că funcția caracteristică a acestei repartiții este

$$((\det \delta^2)(\det \mathbf{A}))^{-\left(m+\frac{k+1}{2}\right)}, \tag{13.7.10}$$

unde $\mathbf{A}$ este matricea definită mai sus prin relațiile (13.7.8). Comparînd acum (13.7.8) și (13.7.10), rezultă că densitatea de probabilitate a lui $\widetilde{\mathbf{Y}}$ se obține din (13.7.9) punînd $m+\frac{k+1}{2}=\frac{\nu}{2}$, adică $m=\frac{\nu-k-1}{2}$ și este deci

$$\frac{(\det \delta^{-2})^{\frac{\nu}{2}} (\det \mathbf{Y})^{\frac{\nu-k-1}{2}} \exp\left(-\frac{1}{2}\operatorname{ur}(\delta^{-2}\mathbf{Y})\right)}{\pi^{\frac{k(k-1)}{4}} 2^{\frac{k\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)\ldots\Gamma\left(\frac{\nu-k+1}{2}\right)} \tag{13.7.11}$$

pentru $\mathbf{Y} \geqslant 0$ și nulă în rest, ceea ce arată că $\widetilde{\mathbf{Y}} \in W(\delta^{-2}, \nu)$ și teorema este deci demonstrată.

Observație. Dacă $\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_\nu$ sînt vectori aleatori independenți din clasa $N(\mathbf{m}_i, \delta^2)$, $1 \leqslant i \leqslant \nu$, atunci $\widetilde{\mathbf{x}}_1 - \mathbf{m}_1, \ldots, \widetilde{\mathbf{x}}_\nu - \mathbf{m}_\nu$ sînt independenți și identic repartizați din clasa $N(0, \delta^2)$ și deci matricea aleatoare

$$\widetilde{\mathbf{Y}} = \sum_{i=1}^{\nu} (\widetilde{\mathbf{x}}_i - \mathbf{m}_i)(\widetilde{\mathbf{x}}_i - \mathbf{m}_i)'$$

aparține de asemenea clasei $W(\delta^{-2}, \nu)$.

Scanned with OKEN Scanner

În cele ce urmează vom prezenta fără demonstrație * principalele proprietăți ale repartiției Wishart.

Dacă $\widetilde{\mathbf{X}} \in W(\psi, \nu)$, atunci

$$M(\widetilde{\mathbf{X}}) = \nu \psi^{-1} \tag{13.7.12}$$

și

$$\operatorname{cov}(\widetilde{x}_{ij}, \widetilde{x}_{rs}) = \frac{1}{\nu - k + 1} (\psi_{ij}^{-1}\psi_{js}^{-1} + \psi_{is}^{-1}\psi_{jr}^{-1}), \tag{13.7.13}$$

unde ψ_{ij}^{-1}, $1 \leqslant i, j \leqslant k$, sînt elementele matricei ψ^{-1}.

Dacă se consideră $\widetilde{\mathbf{X}} \in W(\psi, \nu)$ și se partiționează $\widetilde{\mathbf{X}}$ în modul următor

$$\widetilde{\mathbf{X}} = \begin{pmatrix} \widetilde{\mathbf{X}}_{11} & \widetilde{\mathbf{X}}_{12} \\ \widetilde{\mathbf{X}}_{21} & \widetilde{\mathbf{X}}_{22} \end{pmatrix} \tag{13.7.14}$$

unde $\widetilde{\mathbf{X}}_{11}$ este o matrice pătrată de ordinul r, $1 \leqslant r < k$, se pune problema determinării repartiției marginale a lui $\widetilde{\mathbf{X}}_{11}$. Dacă se partiționează corespunzător ψ și ψ^{-1} în

$$\psi = \begin{pmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{pmatrix}, \quad \psi^{-1} = \begin{pmatrix} \psi_{11}^{-1} & \psi_{12}^{-1} \\ \psi_{21}^{-1} & \psi_{22}^{-1} \end{pmatrix} \tag{13.7.15}$$

atunci se poate arăta că $\widetilde{\mathbf{X}}_{11} \in W(\psi_0, \nu)$, unde

$$\psi_0 = \psi_{11} - \psi_{12} (\psi_{22}^{-1}) \psi_{21}. \tag{13.7.16}$$

O proprietate utilă în aplicații este dată în teorema care urmează

Teorema 13.7.2. *Fie* $\widetilde{\mathbf{x}} \in \mathrm{N}(0, \delta^2)$ *și fie* $\nu\widetilde{\mathbf{X}} \in \mathrm{W}(\delta^{-2}, \nu)$, *unde* $\widetilde{\mathbf{x}}$ *și* $\widetilde{\mathbf{X}}$ *sînt independente. În aceste condiții variabila aleatoare*

$$\widetilde{y} = \frac{\nu - k + 1}{\nu k} \widetilde{\mathbf{x}}' \widetilde{\mathbf{X}}^{-1} \widetilde{\mathbf{x}}$$

aparține clasei $\mathrm{F}(k, \nu - k + 1)$.

---

* Pentru demonstrație se poate vedea Anderson, T. W. (1958).

Scanned with OKEN Scanner

O repartiție strîns legată de repartiția Wishart este *repartiția Wishart inversă*. Spunem că o matrice aleatoare simetrică $\widetilde{\mathbf{X}}$ de ordinul k are *repartiție Wishart inversă de parametrii* ψ *și* ν, unde ψ este o matrice de ordinul k simetrică și pozitiv definită, iar $\nu > 0$, dacă densitatea de probabilitate a lui $\widetilde{\mathbf{X}}$ este de forma

$$f_{Wi}(\mathbf{X} \mid \psi, \nu) = \begin{cases} \dfrac{(\det \psi)^{\frac{\nu}{2}} (\det \mathbf{X}^{-1})^{\frac{\nu+k+1}{2}} \exp\left(-\dfrac{1}{2} \operatorname{ur}(\mathbf{X}^{-1}\psi)\right)}{\pi^{\frac{k(k-1)}{4}} 2^{\frac{k\nu}{2}} \Gamma\left(\dfrac{\nu}{2}\right) \ldots \Gamma\left(\dfrac{\nu-k+1}{2}\right)} \\ \qquad \text{dacă } \mathbf{X} \text{ este pozitiv definită;} \\ 0 \qquad \text{în rest.} \end{cases} \tag{13.7.17}$$

Mulțimea matricelor aleatoare de ordinul k simetrice care au această repartiție va fi notată prin $Wi(\psi, \nu)$.

Legătura dintre repartiția Wishart și repartiția Wishart inversă este precizată în teorema următoare.

Teorema 13.7.3. *Condiția necesară și suficientă ca matricea aleatoare simetrică* $\widetilde{\mathbf{X}}$ *să aparțină clasei* $\mathrm{W}(\psi, \nu)$ *este ca matricea aleatoare simetrică* $\widetilde{\mathbf{Y}} = \widetilde{\mathbf{X}}^{-1}$ *să aparțină clasei* $\mathrm{Wi}(\psi, \nu)$.

Demonstrație. Pentru a arăta necesitatea, fie $\widetilde{\mathbf{X}} \in W(\psi, \nu)$, a cărei densitate de probabilitate este dată de (13.7.1). Considerînd schimbarea de variabile $\mathbf{X} = \mathbf{Y}^{-1}$, al cărui jacobian este egal cu $\det(\mathbf{Y}^{-1})^{k+1}$, rezultă că densitatea de probabilitate $f(\mathbf{Y})$ a lui $\widetilde{\mathbf{Y}}$ se poate scrie astfel

$$f(\mathbf{Y}) = f_W(\mathbf{Y}^{-1} \mid \psi, \nu)\,(\det(\mathbf{Y}^{-1}))^{k+1} =$$
$$= f_W(\mathbf{Y}^{-1} \mid \psi, \nu)(\det \mathbf{Y})^{-(k+1)} = f_{Wi}(\mathbf{Y} \mid \psi, \nu),$$

de unde rezultă că $\widetilde{\mathbf{Y}} \in Wi(\psi, \nu)$.

Suficiența se demonstrează în același mod.

Se poate arăta imediat că repartiția Wishart inversă este o generalizare a repartiției gama inversă. Mai exact, în cazul $k = 1$, avem $Wi(\psi, \nu) = \gamma i(\psi/\nu, \nu)$.

Este de asemenea ușor de văzut că dacă $\widetilde{\mathbf{X}} \in Wi(\psi, \nu)$ și $a > 0$, atunci $a\widetilde{\mathbf{X}} \in Wi(a\psi, \nu)$.

Scanned with OKEN Scanner

Prezentăm în continuare, fără demonstrație *, principalele rezultate privind repartiția Wishart inversă.

Dacă $\widetilde{\mathbf{X}} \in Wi(\psi, \nu)$, atunci

$$M(\widetilde{\mathbf{X}}) = \nu\psi^{-1} \tag{13.7.18}$$

și

$$\operatorname{cov}(\widetilde{x}_{ij}, \widetilde{x}_{rs}) = \frac{1}{\nu - k + 1}(\psi_{ir}^{-1}\psi_{js}^{-1} + \psi_{is}^{-1}\psi_{jr}^{-1}) \tag{13.7.19}$$

cu condiția ca $\nu+1>k$.

Dacă $\widetilde{\mathbf{X}} \in Wi(\psi, \nu)$ și $\widetilde{\mathbf{X}}$ se partiționează ca în (13.7.14) iar ψ ca în (13.7.15), atunci repartiția marginală a lui $\widetilde{\mathbf{X}}_{11}$ este $Wi(\psi_{11}, \nu)$, iar matricea aleatoare

$$\widetilde{\mathbf{X}}_{11|2} = \widetilde{\mathbf{X}}_{11} - \widetilde{\mathbf{X}}_{12}(\widetilde{\mathbf{X}}_{22})^{-1}\,\widetilde{\mathbf{X}}_{21} \tag{13.7.20}$$

aparține clasei $Wi(\psi_0, \nu)$, unde ψ_0 este dat de relația (13.7.16).

13.8. REPARTIȚIA BETA-2 INVERSĂ MULTIDIMENSIONALĂ

Spunem că o matrice aleatoare de ordinul k simetrică $\widetilde{\mathbf{X}}$ are *repartiție beta-2 inversă de ordinul k cu parametrii*, φ, ρ, ν *și* τ, unde φ este o matrice de ordinul k simetrică și pozitiv definită, iar ρ, ν și τ sînt numere reale pozitive, dacă densitatea de probabilitate a lui $\widetilde{\mathbf{X}}$ este de forma

$$f_{\beta 2i}(\mathbf{X} \mid \varphi, \rho, \nu, \tau) = \begin{cases} \dfrac{(\det \tau\varphi)^{\frac{\rho}{2}}\,(\det \mathbf{X})^{\frac{\nu-k-1}{2}}\,[\det(\mathbf{X}+\tau\varphi)]^{-\frac{\nu+\rho}{2}}}{\pi^{\frac{k(k-1)}{4}}\,B_k'\left(\frac{\nu}{2}, \frac{\rho}{2}\right)} \\ \text{dacă } \mathbf{X} \text{ este pozitiv definită,} \\ 0 \qquad \text{în rest,} \end{cases} \tag{13.8.1}$$

* Pentru demonstrații se poate vedea Evans, I.G. (1965) și Tiao, G. C., Zellner, A. (1964).

unde $B'_k(\nu/2,\ \rho/2)$ este dat de relația

$$B'_k\left(\frac{\nu}{2},\ \frac{\rho}{2}\right)^{-1}=\frac{\Gamma\left(\frac{\nu}{2}\right)\ldots\Gamma\left(\frac{\nu-k+1}{2}\right)\Gamma\left(\frac{\rho}{2}\right)\ldots\Gamma\left(\frac{\rho-k+1}{2}\right)}{\Gamma\left(\frac{\nu+\rho}{2}\right)\Gamma\left(\frac{\nu+\rho-1}{2}\right)\ldots\Gamma\left(\frac{\nu+\rho-k+1}{2}\right)}. \tag{13.8.2}$$

Mulțimea matricelor aleatoare cu această repartiție va fi notată prin $\beta 2i(\varphi,\ \rho,\ \nu,\ \tau)$. Se observă ușor că $\beta 2i(\varphi,\ \rho,\ \nu,\ \tau) = \beta\, 2i(\tau\varphi,\ \rho,\ \nu,\ 1)$, ceea ce arată că parametrul τ nu este esențial.

De asemenea se arată ușor că această repartiție este o generalizare a repartiției beta-2 inversă unidimensională. Mai precis, în cazul unidimensional avem $\beta\, 2i(\varphi,\ \rho,\ \nu,\ \tau) = \beta\, 2i\left(\frac{\nu}{2},\ \frac{\rho}{2},\ \tau\varphi\right)$.

Se poate observa imediat că dacă $\widetilde{\mathbf{X}} \in \beta\, 2i(\varphi,\ \rho,\ \nu,\ \tau)$, iar $a > 0$, atunci $a\widetilde{\mathbf{X}} \in \beta\, 2i\left(\varphi,\ \rho,\ \nu,\ \frac{\tau}{a}\right)$.

Teorema care urmează arată modul în care apare această repartiție.

Teorema 13.8.1. *Fie* $\widetilde{\mathbf{Y}}$ *o matrice aleatoare cu proprietatea că* $\widetilde{\mathbf{Y}}^{-1}$ *aparține clasei* $Wi(\tau\varphi,\ \rho)$ *și fie* $\widetilde{\mathbf{X}}$ *o altă matrice aleatoare de ordinul* k *cu proprietatea că densitatea de probabilitate a lui* $\widetilde{\mathbf{X}}$ *condiționată de* $\widetilde{\mathbf{Y}} = \mathbf{Y}$ *este* $f_W(\mathbf{X} \mid \mathbf{Y}^{-1},\ \nu)$.

În aceste condiții rezultă că $\widetilde{\mathbf{X}}$ *aparține clasei* $\beta\, 2i(\varphi,\ \rho,\ \nu,\ \tau)$.

Demonstrație. Densitatea de probabilitate $f(\mathbf{X},\ \mathbf{Y})$ a lui $(\widetilde{\mathbf{X}},\ \widetilde{\mathbf{Y}})$ este dată evident de relația

$$f(\mathbf{X},\ \mathbf{Y}) = f_{Wi}(\mathbf{Y} \mid \tau\varphi,\ \rho)\, f_W(\mathbf{X} \mid \mathbf{Y}^{-1},\ \nu).$$

Aplicînd relațiile (13.7.1) și (13.7.17) rezultă

$$f(\mathbf{X},\ \mathbf{Y}) =$$

$$= \frac{(\det \tau\varphi)^{\frac{\rho}{2}}(\det \mathbf{X})^{\frac{\nu-k-1}{2}}(\det \mathbf{Y}^{-1})^{\frac{\nu+\rho+k+1}{2}}\exp\left[-\frac{1}{2}\mathrm{ur}\left(\mathbf{Y}^{-1}(\mathbf{X}+\tau\varphi)\right)\right]}{\pi^{\frac{k(k-1)}{2}}\, 2^{\frac{k(\nu+\rho)}{2}}\,\Gamma\left(\frac{\nu}{2}\right)\ldots\Gamma\left(\frac{\nu-k+1}{2}\right)\Gamma\left(\frac{\rho}{2}\right)\ldots\Gamma\left(\frac{\rho-k+1}{2}\right)}$$

Scanned with OKEN Scanner

sau încă

$$f(\mathbf{X}, \mathbf{Y}) = \frac{(\det \tau\varphi)^{\frac{\rho}{2}} (\det \mathbf{X})^{\frac{\nu-k-1}{2}} [\det (\mathbf{X}+\tau\varphi)]^{-\frac{\nu+\rho}{2}}}{\pi^{\frac{k(k-1)}{4}} B_k'\left(\frac{\nu}{2}, \frac{\rho}{2}\right)} \cdot$$

$$\cdot \frac{(\det \mathbf{Y}^{-1})^{\frac{\nu+\rho+k+1}{2}} [\det (\mathbf{X}+\tau\varphi)]^{\frac{\nu+\rho}{2}} \exp\left[-\frac{1}{2} \operatorname{ur} (\mathbf{X}+\tau\varphi)\right]}{\pi^{\frac{k(k-1)}{4}} 2^{\frac{k(\nu+\rho)}{2}} \Gamma\left(\frac{\nu+\rho}{2}\right) \ldots \Gamma\left(\frac{\nu+\rho-k+1}{2}\right)}.$$

Ținînd seama că a doua fracție este chiar $f_{Wi}(\mathbf{Y} \mid \mathbf{X}+\tau\varphi, \nu+\rho)$ și că prima fracție este independentă de $\mathbf{Y}$, rezultă că integrînd $f(\mathbf{X}, \mathbf{Y})$ în raport cu $\mathbf{Y}$, valoarea integralei este egală cu prima fracție și prin urmare densitatea de probabilitate a lui $\widetilde{\mathbf{X}}$ este $f_{\beta 2i}(\varphi, \rho, \nu, \tau)$, ceea ce demonstrează teorema.

Să observăm că, pentru $\widetilde{\mathbf{X}} \in \beta 2i(\varphi, \rho, \nu, \tau)$, putem scrie

$$M(\widetilde{\mathbf{X}}) = M(M(\widetilde{\mathbf{X}} \mid \widetilde{\mathbf{Y}}^{-1})) = M(\nu\widetilde{\mathbf{Y}}) = \frac{\nu\tau}{\rho-k-1}\varphi, \quad (13.8.2)$$

cu condiția ca $\rho > k+1$.

Tot pe aceeași cale se obține repartiția marginală a lui $\widetilde{\mathbf{X}}_{11}$ dacă $\widetilde{\mathbf{X}} \in \beta\, 2i(\varphi, \rho, \nu, \tau)$, unde $\widetilde{\mathbf{X}}$ este partiționat ca în (13.7.4). Se poate vedea ușor că dacă φ este partiționată corespunzător în

$$\varphi = \begin{pmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{pmatrix},$$

atunci $\widetilde{\mathbf{X}}_{11} \in \beta\, 2i(\varphi_{11}, \rho, \nu, \tau)$.

13.9. REPARTIȚIILE STATISTICILOR ASOCIATE UNUI PROCES NORMAL DE GENERARE DE DATE

În cele ce urmează ne vom ocupa cu determinarea repartițiilor statisticilor care sînt asociate în mod natural unui proces de generare de date normal (vezi § 11.6 și § 11.7), precum și a unor statistici utile construite cu ajutorul acestora.

13.9.1. CAZUL UNIDIMENSIONAL

Pe baza rezultatelor stabilite în acest capitol putem determina repartițiile asociate statisticilor $\tilde{\bar{x}}$, $\tilde{s}^2$ și $\tilde{s}^2(m)$ introduse în § 11.6.1 pentru un proces normal unidimensional de generare de date $(\tilde{x}_i)$ de parametrii m și $\sigma^2 > 0$, precum și pentru statistici utile construite cu ajutorul acestora.

Teorema 13.9.1. *Dacă* $\tilde{x}_1, \ldots, \tilde{x}_n$ *constituie o selecție de volum* n *din populația variabilei aleatoare* $\tilde{x} \in N(m, \sigma^2)$, *atunci:*

(a) Media de selecție $\tilde{\bar{x}}$ *aparține clasei* $N(m, \sigma^2/n)$;

(b) Statistica

$$\tilde{q}_1 = \frac{n(\tilde{\bar{x}} - m)^2}{\sigma^2} \tag{13.9.1}$$

aparține clasei $\chi^2(1)$;

(c) Statistica

$$\tilde{q}_2 = \frac{(n-1)\tilde{s}^2}{\sigma^2} \tag{13.9.2}$$

aparține clasei $\chi^2(n-1)$ *și este independentă de media de selecție* $\tilde{\bar{x}}$;

(d) Statistica $\tilde{s}^2$ *aparține clasei* $\gamma 2(1/\sigma^2, n-1)$ *sau, echivalent, clasei* $\chi^2(n-1, \sigma^2/(n-1))$ *și este independentă de media de selecție* $\tilde{\bar{x}}$;

(e) Statistica

$$\tilde{q}_3 = \frac{n}{\sigma^2}\tilde{s}^2(m) \tag{13.9.3}$$

aparține clasei $\chi^2(n)$ *și este independentă de media de selecție* $\tilde{\bar{x}}$. *Mai mult,* $\tilde{s}^2(m)$ *aparține clasei* $\gamma 2(1/\sigma^2, n)$ *sau, echivalent, clasei* $\chi^2(n, \sigma^2/n)$;

(f) Statistica

$$\tilde{\tau} = \frac{(\tilde{\bar{x}} - m)}{\sqrt{\tilde{s}^2/n}} \tag{13.9.4}$$

aparține clasei $S^*(n-1)$.

Demonstrație. (*a*) Afirmația este un caz particular al proprietății unei combinații liniare de variabile aleatoare normale independente.

(*b*) Afirmația rezultă din (*a*) și din teorema 13.2.1.

(*c*) Să considerăm matricea de ordinul n

$$\mathbf{D} = \begin{pmatrix} 1/\sqrt{n} & 1/\sqrt{n} & 1/\sqrt{n} & 1/\sqrt{n} & \dots & 1/\sqrt{n} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 & 0 & \dots & 0 \\ 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1/\sqrt{n(n-1)} & 1/\sqrt{n(n-1)} & \dots & \dots & \dots & (1-n)/\sqrt{n(n-1)} \end{pmatrix}$$

Se constată ușor că această matrice este ortonormală. Dacă se consideră variabilele aleatoare independente și identic repartizate

$$\tilde{y}_i = \frac{\tilde{x}_i - m}{\sigma}, \qquad 1 \leqslant i \leqslant n,$$

care aparțin clasei $N(0, 1)$ și dacă notăm prin $\tilde{\mathbf{y}}$ vectorul aleator $(\tilde{y}_1, \dots, \tilde{y}_n)'$, atunci rezultă că vectorul aleator $\tilde{\mathbf{z}} = \mathbf{D}\tilde{\mathbf{y}}$ aparține clasei $N(O, I)$, iar componentele sale $\tilde{z}_1, \dots, \tilde{z}_n$ sînt variabile aleatoare independente și identic repartizate din clasa $N(0, 1)$, conform cu observația de la teorema 10.3.2.

Scanned with OKEN Scanner

Este clar că $\tilde{z}_1, \tilde{z}_2^2, \ldots, \tilde{z}_n^2$ sînt de asemenea variabile aleatoare independente. Este limpede că avem

$$\tilde{z}_1 = \frac{\tilde{y}_1 + \ldots + \tilde{y}_n}{\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{(\tilde{x}_i - m)}{\sigma} = \frac{\tilde{\bar{x}} - m}{\sigma\sqrt{n}} \quad (13.9.5)$$

și

$$\sum_{i=2}^{n} \tilde{z}_i^2 = \tilde{\mathbf{z}}'\tilde{\mathbf{z}} - \tilde{z}_1^2 = (\mathbf{D}\tilde{\mathbf{y}})'(\mathbf{D}\tilde{\mathbf{y}}) - \tilde{z}_1^2 = \tilde{\mathbf{y}}'\mathbf{D}'\mathbf{D}\tilde{\mathbf{y}} - \tilde{z}_1^2 = \sum_{i=1}^{n} \tilde{y}_i^2 - \tilde{z}_1^2$$

sau

$$\sum_{i=2}^{n} \tilde{z}_i^2 = \sum_{i=1}^{n} \frac{(\tilde{x}_i - m)^2}{\sigma^2} - \left(\frac{\tilde{\bar{x}} - m}{\sigma/\sqrt{n}} \right)^2. \quad (13.9.6)$$

Observînd că

$$\tilde{q}_2 = \sum_{i=1}^{n} \left(\frac{\tilde{x}_i - \tilde{\bar{x}}}{\sigma} \right)^2 = \sum_{i=1}^{n} \left(\frac{\tilde{x}_i - m}{\sigma} + \frac{m - \tilde{\bar{x}}}{\sigma} \right)^2$$

sau

$$\tilde{q}_2 = \sum_{i=1}^{n} \left(\frac{\tilde{x}_i - m}{\sigma} \right)^2 - \left(\frac{\tilde{\bar{x}} - m}{\sigma/\sqrt{n}} \right)^2, \quad (13.9.7)$$

din (13.9.5), (13.9.6) și (13.9.7) urmează că variabilele $\tilde{z}_1$ și $\tilde{q}_2$ sînt independente, ceea ce implică faptul că $\tilde{\bar{x}}$ și $\tilde{q}_2$ sînt independente.

Din relațiile (13.9.6) și (13.9.7) se deduce imediat că $\tilde{q}_2$ aparține clasei $\chi^2(n-1)$. Într-adevăr, avem

$$\tilde{q}_2 = \sum_{i=2}^{n} \tilde{z}_i^2$$

și $\tilde{z}_i \in N(0, 1)$, $2 \leqslant i \leqslant n$, și deci $\tilde{q}_2 \in \chi^2(n-1)$, conform teoremei 13.2.1.

(d) În general, dacă $\tilde{y}$ aparține clasei $\chi^2(\nu)$ și $\alpha > 0$, atunci $\alpha\tilde{y}$ aparține clasei $\gamma 2(1/\alpha\nu, \nu)$, cum rezultă imediat din relațiile

(13.2.3) și (13.2.4). Deoarece $\tilde{q}_2 \in \chi^2(n-1)$ și $\sigma^2/(n-1) > 0$, rezultă că

$$\tilde{s}^2 = \frac{\sigma^2}{n-1}\tilde{q}_2$$

aparține clasei $\gamma 2(1/\sigma^2, n-1)$ sau, echivalent, clasei $\chi^2(n-1, \sigma^2/(n-1))$, cum rezultă din (13.2.5). Afirmația privind independența statisticilor $\tilde{\tilde{x}}$ și $\tilde{s}^2$ rezultă evident din independența variabilelor aleatoare $\tilde{\tilde{x}}$ și $\tilde{q}_2$.

(e) Deoarece

$$\tilde{q}_3 = \frac{n}{\sigma^2} \cdot \frac{1}{n}\sum_{i=1}^{n} (\tilde{x}_i - m)^2 = \sum_{i=1}^{n} \left(\frac{\tilde{x}_i - m}{\sigma}\right)^2,$$

conform teoremei 13.2.1 rezultă că $\tilde{q}_3 \in \chi^2(n)$. Faptul că $\tilde{q}_3$ este independent de $\tilde{\tilde{x}}$ se demonstrează pe o cale analoagă celei utilizate în demonstrația independenței variabilelor aleatoare $\tilde{\tilde{x}}$ și $\tilde{q}_2$. Celelalte afirmații se deduc din observația generală de la punctul *(d)*.

(f) Deoarece

$$\tilde{t} = \frac{(\tilde{\tilde{x}} - m)}{\sqrt{\frac{\sigma^2 \tilde{q}_2}{n(n-1)}}} = \frac{\frac{\tilde{\tilde{x}} - m}{\sigma/\sqrt{n}}}{\sqrt{\frac{\tilde{q}_2}{n-1}}},$$

afirmația decurge imediat din teorema 13.4.2, ținînd totodată seama de faptul că din *(c)* rezultă că

$$\frac{\tilde{\tilde{x}} - m}{\sigma\sqrt{n}} \in N(0, 1), \quad \tilde{q}_2 \in \chi^2(n-1)$$

sînt variabile aleatoare independente.

Pentru compararea a două procese normale de generare de date se utilizează o serie de statistici, care sînt de asemenea utile în studiul procesului normal multidimensional.

Fie $\tilde{x}_{11}, \ldots, \tilde{x}_{1n_1}$ o selecție de volum n_1 din $\tilde{x}_1 \in N(m_1, \sigma_1^2)$ și $\tilde{x}_{21}, \ldots, \tilde{x}_{2n_2}$ o selecție de volum n_2 din $\tilde{x}_2 \in N(m_2, \sigma_2^2)$. Vom presupune ca aceste selecții sînt *independente* în sensul că variabilele aleatoare $\tilde{x}_{11}, \ldots, \tilde{x}_{1n_1}$, $\tilde{x}_{21}, \ldots, \tilde{x}_{2n_2}$ sînt independente. În cele ce urmează vom utiliza următoarele statistici:

$$\tilde{\bar{x}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \tilde{x}_{ij}, \quad i = 1, 2, \tag{13.9.8}$$

$$\tilde{s}_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (\tilde{x}_{ij} - \tilde{\bar{x}}_i)^2, \quad i = 1, 2, \tag{13.9.9}$$

$$\tilde{s}_i^2(m_i) = \frac{1}{n_i} \sum_{j=1}^{n_i} (\tilde{x}_{ij} - m_i)^2, \quad i = 1, 2, \tag{13.9.10}$$

$$\tilde{s}^2 = \frac{(n_1 - 1)\tilde{s}_1^2 + (n_2 - 1)\tilde{s}_2^2}{n_1 + n_2 - 2}. \tag{13.9.11}$$

Teorema 13.9.2. *Cu notațiile și ipotezele specificate mai sus avem:*

(a) Statistica

$$\tilde{\delta}_1 = \frac{(\tilde{\bar{x}}_1 - \tilde{\bar{x}}_2) - (m_1 - m_2)}{(\sigma_1^2/n_1 + \sigma_2^2/n_2)^{1/2}} \tag{13.9.12}$$

aparține clasei $N(0, 1)$;

(b) Statistica

$$\tilde{f}_1 = \frac{\tilde{s}_1^2/\sigma_1^2}{\tilde{s}_2^2/\sigma_2^2} \tag{13.9.13}$$

aparține clasei $F(n_1 - 1, n_2 - 1)$;

(c) Statistica

$$\tilde{f}_2 = \frac{\tilde{s}_1^2(m_1)/\sigma_1^2}{\tilde{s}_2^2(m_2)/\sigma_2^2} \tag{13.9.14}$$

aparține clasei $F(n_1, n_2)$;

(d) Dacă, în plus, $\sigma_1^2 = \sigma_2^2 = \sigma^2$, *atunci statistica*

$$\tilde{\delta}_2 = \frac{(\tilde{\bar{x}}_1 - \tilde{\bar{x}}_2) - (m_1 - m_2)}{\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\tilde{s}^2\right]^{1/2}} \tag{13.9.15}$$

aparține clasei $S^*(n_1 + n_2 - 2)$.

Demonstrație. (a) Din teorema precedentă rezultă că $\tilde{\bar{x}}_i \in N(m_i, \sigma_i^2/n_i)$, pentru $i = 1, 2$. În plus, $\tilde{\bar{x}}_1$ și $\tilde{\bar{x}}_2$ sînt variabile aleatoare independente. Rezultă atunci că $\tilde{\bar{x}}_1 - \tilde{\bar{x}}_2$ aparține clasei $N(m_1 - m_2, \sigma_1^2/n_1 + \sigma_2^2/n_2)$, de unde se deduce imediat că $\tilde{\delta}_1 \in N(0, 1)$.

(b) Din teorema precedentă rezultă că variabilele aleatoare

$$\tilde{h}_i = \frac{(n_i - 1)\tilde{s}_i^2}{\sigma_i^2} \in \chi^2(n_i - 1), \quad i = 1, 2, \tag{13.9.16}$$

și, în plus, $\tilde{h}_1$ și $\tilde{h}_2$ sînt independente. Din teorema 13.3.3 rezultă atunci că raportul

$$\frac{\tilde{h}_1/(n_1 - 1)}{\tilde{h}_2/(n_2 - 1)} = \frac{\tilde{s}_1^2/\sigma_1^2}{\tilde{s}_2^2/\sigma_2^2} \tag{13.9.17}$$

aparține clasei $F(n_1 - 1, n_2 - 1)$.

(c) Din teorema precedentă rezultă că variabilele aleatoare

$$\tilde{h}_i(m_i) = \frac{n_i \tilde{s}_i^2(m_i)}{\sigma_i^2}, \quad i = 1, 2, \tag{13.9.18}$$

aparțin clasei $\chi^2(n_i)$. În plus, $\tilde{h}_1(m_1)$ și $\tilde{h}_2(m_2)$ sînt variabile aleatoare independente. Utilizînd din nou teorema 13.3.3, rezultă că raportul

$$\frac{\tilde{h}_1(m_1)/n_1}{\tilde{h}_2(m_2)/n_2} = \frac{\tilde{s}_1^2(m_1)}{\tilde{s}_2^2(m_2)} \tag{13.9.19}$$

aparține clasei $F(n_1, n_2)$.

Scanned with OKEN Scanner

(d) Utilizînd variabilele aleatoare $\widetilde{h}_1$ și $\widetilde{h}_2$ date de (13.9.16) putem scrie

$$\widetilde{s}^2 = \frac{\sigma_1^2\widetilde{h}_1 + \sigma_2^2\widetilde{h}_2}{n_1 + n_2 - 2} = \frac{\sigma^2(\widetilde{h}_1 + \widetilde{h}_2)}{n_1 + n_2 - 2}$$

și deci $\widetilde{\delta}_2$ dat de (13.9.15) devine

$$\widetilde{\delta}_2 = \frac{(\widetilde{\bar{x}}_1 - \widetilde{\bar{x}}_2) - (m_1 - m_2)}{\left(\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}\right)^{1/2} \sqrt{\frac{\widetilde{h}_1 + \widetilde{h}_2}{n_1 + n_2 - 2}}} = \frac{\widetilde{\delta}_1}{\sqrt{\frac{\widetilde{h}_1 + \widetilde{h}_2}{n_1 + n_2 - 2}}}.$$

Cum $\widetilde{h}_1 \in \chi^2(n_1 - 1)$ și $\widetilde{h}_2 \in \chi^2(n_2 - 1)$ sînt variabile aleatoare independente, rezultă că $\widetilde{h}_1 + \widetilde{h}_2 \in \chi^2(n_1 + n_2 - 2)$. Din afirmația *(a)* a teoremei noastre rezultă că $\widetilde{\delta}_1 \in N(0, 1)$. Deoarece $\widetilde{\delta}_1$ și $\widetilde{h}_1 + \widetilde{h}_2$ sînt variabile aleatoare independente, obținem imediat $\widetilde{\delta}_2 \in S^*(n_1 + n_2 - 2)$, conform teoremei 13.4.2, și teorema este demonstrată.

13.9.2. CAZUL MULTIDIMENSIONAL

Ca și în cazul unidimensional, vom determina repartițiile asociate statisticilor uzuale introduse în § 11.7.1 pentru un proces k—dimensional $(\widetilde{\mathbf{x}}_i)$ normal de parametrii $\mathbf{m}$ și $\boldsymbol{\delta}^2$, precum și pentru statisticile utile construite cu ajutorul acestora.

Teorema 13.9.3. *Dacă* $\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_n$ *constituie o selecție de volum* n *din populația vectorului aleator* $\widetilde{\mathbf{x}}$ *din clasa* $\mathrm{N}(\mathbf{m}, \boldsymbol{\delta}^2)$, *atunci:*

(a) Media de selecție $\widetilde{\bar{\mathbf{x}}}$ *aparține clasei* $\mathrm{N}\left(\mathbf{m}, \frac{1}{\mathrm{n}}\boldsymbol{\delta}^2\right)$;

(b) Statistica

$$\widetilde{\mathbf{w}}_1 = (\mathrm{n}-1)\widetilde{\mathbf{s}}^2 \tag{13.9.20}$$

aparține clasei $\mathrm{W}(\boldsymbol{\delta}^{-2}, \mathrm{n}-1)$, *statistica* $\widetilde{\mathbf{s}}^2$ *aparține clasei* $\mathrm{W}((\mathrm{n}-1)\boldsymbol{\delta}^{-2}, \mathrm{n}-1)$, *iar* $\widetilde{\mathbf{w}}_1$ *și* $\widetilde{\mathbf{s}}^2$ *sînt, fiecare, independente de* $\widetilde{\bar{\mathbf{x}}}$;

Scanned with OKEN Scanner

(c) Statistica

$$\tilde{w}_2 = n\,\tilde{s}^2(m) \tag{13.9.21}$$

aparține clasei $W(\delta^{-2}, n)$, *iar statistica* $\tilde{s}^2(m)$ *aparține clasei* $W(n\,\delta^{-2}, n)$;

(d) Statistica

$$\tilde{q}_1 = \frac{1}{\sigma_*^2} nk\tilde{s}_*^2(m) \tag{13.9.22}$$

aparține clasei χ^2 (kn), *iar statistica* $\tilde{s}_*^2(m)$ *aparține clasei* $\gamma 2(1/\sigma_*^2, kn)$;

(e) Statistica

$$\tilde{q}_2 = \frac{k(n-1)\tilde{s}_*^2}{\sigma_*^2} \tag{13.9.23}$$

aparține clasei $\chi^2(k(n-1))$, *statistica* $\tilde{s}_*^2$ *aparține clasei* $\gamma 2(1/\sigma_*^2, k(n-1))$, *iar* $\tilde{s}_*^2$ *este independentă de* $\tilde{\bar{x}}$;

(f) Statistica

$$\tilde{q}_3 = n(\tilde{\bar{x}} - m)'\,\delta^{-2}(\tilde{\bar{x}} - m) \tag{13.9.24}$$

aparține clasei $\chi^2(k)$.

Demonstrație. (a), (b). Să considerăm o matrice ortogonală $\mathbf{D} = (d_{ij})$ de ordinul k având prima linie $(1/\sqrt{n}, \ldots, 1/\sqrt{n})$. Vectorii aleatori $\tilde{y}_i$, $1 \leqslant i \leqslant n$, definiți de relația

$$\tilde{y}_i = \sum_{j=1}^{n} d_{ij}\tilde{x}_j, \quad 1 \leqslant i \leqslant n,$$

sînt independenți și aparțin clasei $N\left(\sum_{j=1}^{n} d_{ij}m,\ \delta^2\right)$, conform teoremei 10.3.7. Însă

$$\tilde{y}_1 = \sum_{j=1}^{n} d_{1j}\tilde{x}_j = \sum_{j=1}^{n} \frac{1}{\sqrt{n}}\tilde{x}_j = \frac{1}{\sqrt{n}}\, n\tilde{\bar{x}} = \sqrt{n}\tilde{\bar{x}}$$

Scanned with OKEN Scanner

și

$$\sum_{j=1}^{n} d_{1j}\mathrm{m} = \sqrt{n}\,\mathrm{m},$$

ceea ce arată că $\tilde{\bar{x}} = \frac{1}{\sqrt{n}}\tilde{y}_1$ aparține clasei $N\left(\mathrm{m}, \frac{1}{n}\delta^2\right)$.

Deoarece

$$\tilde{w}_1 = (n-1)\tilde{s}^2 = \sum_{i=1}^{n}(\tilde{x}_i - \tilde{\bar{x}})(\tilde{x}_i - \tilde{\bar{x}})' = \sum_{i=1}^{n}\tilde{x}_i\tilde{x}_i' - n\tilde{\bar{x}}\tilde{\bar{x}}' =$$

$$= \sum_{i=1}^{n}\tilde{y}_i\tilde{y}_i' - \tilde{y}_1\tilde{y}_1' = \sum_{i=2}^{n} y_i y_i',$$

rezultă mai întîi că $\tilde{w}_1$ și $\tilde{s}^2$ sînt, fiecare, independenți de $\tilde{\bar{x}}$ (deoarece $\tilde{y}_2, \ldots, \tilde{y}_n$ sînt independenți de $\tilde{y}_1$). Pe de altă parte, deoarece $\tilde{y}_2, \ldots, \tilde{y}_n$ sînt independenți și normal repartizați, din teorema 13.7.1 rezultă că $\tilde{w}_1 \in W(\delta^{-2}, n-1)$ și deci $\tilde{s}^2 \in W((n-1)\delta^{-2}, n-1)$.

(c) Deoarece

$$\tilde{w}_2 = n\tilde{s}^2(\mathrm{m}) = \sum_{i=1}^{n}(\tilde{x}_i - \mathrm{m})(\tilde{x}_i - \mathrm{m})',$$

iar vectorii aleatori $\tilde{x}_1, \ldots, \tilde{x}_n$ sînt independenți și aparțin clasei $N(\mathrm{m}, \delta^2)$, din teorema 13.7.1 rezultă imediat că $\tilde{w}_2 \in W(\delta^{-2}, n)$ și deci $\tilde{s}^2(\mathrm{m}) \in W(n\delta^{-2}, n)$.

(d) Ținînd seama de relațiile (11.7.2) și (11.7.7) rezultă

$$\tilde{q}_1 = \frac{1}{\sigma_*^2} kn\,\tilde{s}_*^2(\mathrm{m}) = \sum_{i=1}^{n}(\tilde{x}_i - \mathrm{m})'\,\delta^{-2}(\tilde{x}_i - \mathrm{m}).$$

Observînd acum că $\tilde{x}_i$, $1 \leqslant i \leqslant n$, sînt vectori aleatori normali nesingulari din clasa $N(\mathrm{m}, \delta^2)$, din teorema 13.2.3 rezultă că variabilele aleatoare independente

$$(\tilde{x}_i - \mathrm{m})'\,\delta^{-2}(\tilde{x}_i - \mathrm{m}), \qquad 1 \leqslant i \leqslant n,$$

aparțin clasei $\chi^2(k)$ și deci suma lor aparține clasei $\chi^2(kn)$, ceea ce demonstrează prima afirmație.

Deoarece

$$s_*^2(\mathrm{m}) = \frac{\sigma_*^2}{kn}\tilde{q}_1,$$

iar $\tilde{q}_1 \in \chi^2(kn) = \gamma(kn/2,\ 1/2)$, rezultă imediat că $\tilde{s}_*^2(\mathrm{m}) \in \gamma(kn/2,\ kn/2\sigma_*^2)$, ceea ce este echivalent cu $\tilde{s}_*^2(\mathrm{m}) \in \gamma 2(1/\sigma_*^2,\ kn)$.

(c) Ținînd seama de relațiile (11.7.2) și (11.7.7), rezultă că

$$\tilde{q}_2 = \frac{k(n-1)\tilde{s}_*^2}{\sigma_*^2} = \sum_{i=1}^{n}(\tilde{\mathbf{x}}_i - \tilde{\bar{\mathbf{x}}})'\boldsymbol{\delta}^{-2}(\tilde{\mathbf{x}}_i - \tilde{\bar{\mathbf{x}}}).$$

Deoarece $\boldsymbol{\delta}^{-2}$ este simetrică și pozitiv definită, există o matrice nesingulară $\mathbf{D}$ astfel încît $\mathbf{D}'\mathbf{D} = \boldsymbol{\delta}^{-2}$. Vectorii aleatori

$$\tilde{\mathbf{y}}_i = \mathbf{D}(\tilde{\mathbf{x}}_i - \mathrm{m}), \qquad 1 \leqslant i \leqslant n,$$

sînt independenți și identic repartizați din clasa $N(\mathbf{0}, \mathbf{I})$, conform teoremei 10.3.2.

Vectorul aleator $\tilde{\bar{\mathbf{y}}}$ definit de

$$\tilde{\bar{\mathbf{y}}} = \frac{1}{n}\sum_{i=1}^{n}\tilde{\mathbf{y}}_i$$

poate fi exprimat și sub forma

$$\tilde{\bar{\mathbf{y}}} = \mathbf{D}(\tilde{\bar{\mathbf{x}}} - \mathrm{m}).$$

Rezultă atunci că

$$\tilde{\mathbf{y}}_i - \tilde{\bar{\mathbf{y}}} = \mathbf{D}(\tilde{\mathbf{x}}_i - \mathrm{m})$$

și

$$(\tilde{\mathbf{y}}_i - \tilde{\bar{\mathbf{y}}})'(\tilde{\mathbf{y}}_i - \tilde{\bar{\mathbf{y}}}) = (\tilde{\mathbf{x}}_i - \mathrm{m})'\mathbf{D}'\mathbf{D}(\tilde{\mathbf{x}}_i - \mathrm{m}) = (\tilde{\mathbf{x}}_i - \mathrm{m})'\boldsymbol{\delta}^{-2}(\tilde{\mathbf{x}}_i - \mathrm{m}),$$

de unde rezultă

$$\tilde{q}_2 = \sum_{i=1}^{n}(\tilde{\mathbf{y}}_i - \tilde{\bar{\mathbf{y}}})'(\tilde{\mathbf{y}}_i - \tilde{\bar{\mathbf{y}}}) = \sum_{i=1}^{n}\sum_{j=1}^{k}(\tilde{y}_{ij} - \tilde{\bar{y}}_j)^2 =$$

$$= \sum_{j=1}^{k}\sum_{i=1}^{n}(\tilde{y}_{ij} - \tilde{\bar{y}}_j)^2,$$

Scanned with OKEN Scanner

unde $\tilde{y}_{ij}$, $1 \leqslant j \leqslant k$, și $\tilde{\bar{y}}_j$, $1 \leqslant j \leqslant k$, sînt componentele vectorilor aleatori $\tilde{\mathbf{y}}_i$ și $\tilde{\bar{\mathbf{y}}}$, respectiv. Deoarece $\{\tilde{y}_{ij}\}$, $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant k$, sînt variabile aleatoare independente și identic repartizate din clasa $N(0, 1)$, din teorema 13.9.1 rezultă că

$$\sum_{i=1}^{n} (\tilde{y}_{ij} - \tilde{\bar{y}}_j)^2$$

sînt variabile aleatoare din clasa $\chi^2(n-1)$, independente de $\tilde{\bar{\mathbf{y}}}$ și deci de $\tilde{\bar{\mathbf{x}}}$. Rezultă atunci că $\tilde{q}_2 \in \chi^2(k(n-1))$ și $\tilde{q}_2$ este independentă de $\tilde{\bar{\mathbf{x}}}$.

În sfîrșit, faptul că $\tilde{s}_*^2 \in \gamma 2(1/\sigma_*^2,\ k(n-1))$ și că $\tilde{s}_*^2$ este independentă de $\tilde{\bar{\mathbf{x}}}$ se deduce imediat din $\tilde{q}_2 \in \chi^2(k(n-1))$ și din faptul că $\tilde{q}_2$ este independentă de $\tilde{\bar{\mathbf{x}}}$.

(f) Rezultatul decurge imediat din teorema 13.2.3 și din punctul *(a)* al teoremei noastre.

Teorema 13.9.4. *Dacă* $\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_n$ *este o selecție de volum* n *din populația vectorului aleator* $\tilde{\mathbf{x}} \in \mathrm{N}(\mathrm{m}, \delta^2)$, *atunci:*

(a) Dacă $\mathrm{n} \geqslant \mathrm{k}$, *atunci statistica*

$$\tilde{\mathrm{f}}_1 = \frac{\mathrm{n}(\mathrm{n}-\mathrm{k})}{\mathrm{k}(\mathrm{n}-1)} (\tilde{\bar{\mathbf{x}}} - \mathrm{m})' \tilde{\mathrm{s}}^{-2} (\tilde{\bar{\mathbf{x}}} - \mathrm{m}) \qquad (13.9.25)$$

aparține clasei $\mathrm{F}(\mathrm{k}, \mathrm{n}-\mathrm{k})$, *unde* $\tilde{\mathrm{s}}^{-2} = (\tilde{\mathrm{s}}^2)^{-1}$;

(b) Dacă $\mathrm{n} \geqslant 2$, *atunci statistica*

$$\tilde{\mathrm{f}}_2 = \frac{\mathrm{n}}{\mathrm{k}} (\tilde{\bar{\mathbf{x}}} - \mathrm{m})' \tilde{\mathrm{s}}_*^{-2} \delta_*^{-2} (\tilde{\bar{\mathbf{x}}} - \mathrm{m}) \qquad (13.9.26)$$

aparține clasei $\mathrm{F}(\mathrm{k}, \mathrm{k}(\mathrm{n}-1))$.

Demonstrație. (a) Din teorema precedentă rezultă că $\sqrt{n}(\tilde{\bar{\mathbf{x}}} - \mathrm{m}) \in N(\mathbf{0}, \delta^2)$ și $\tilde{\mathrm{s}}^2 \in W((n-1)\,\delta^{-2}, n-1)$, adică $(n-1)\tilde{\mathrm{s}}^2$ aparține clasei $W(\delta^{-2}, n-1)$. Din teorema 13.7.2 rezultă atunci că variabila aleatoare

$$\frac{n-k}{k(n-1)} (\sqrt{n}(\tilde{\bar{\mathbf{x}}} - \mathrm{m}))' (\tilde{\mathrm{s}}^2)^{-1} (\sqrt{n}(\tilde{\bar{\mathbf{x}}} - \mathrm{m})) = \tilde{f}_1$$

aparține clasei $F(k, n-k)$.

Scanned with OKEN Scanner

(b) Din teorema precedentă urmează că statisticile $\tilde{q}_2$ și $\tilde{q}_3$ date de relațiile (13.9.23) și (13.9.24) aparțin respectiv claselor $\chi^2(k(n-1))$ și $\chi^2(k)$. În plus, $\tilde{q}_2$ și $\tilde{q}_3$ sînt independente deoarece $\tilde{\tilde{\mathbf{x}}}$ și $\tilde{s}_*^2$ sînt independente, conform afirmației *(e)* de la teorema precedentă. Din teorema 13.3.3 rezultă atunci că raportul

$$\frac{\tilde{q}_3/k}{\tilde{q}_2/k(n-1)} = \frac{k(n-1)}{k} \cdot \frac{n(\tilde{\tilde{\mathbf{x}}} - \mathbf{m})' \, \boldsymbol{\delta}^{-2}(\tilde{\tilde{\mathbf{x}}} - \mathbf{m})}{k(n-1)\tilde{s}_*^2/\sigma^2} =$$

$$= \frac{n}{k} (\tilde{\tilde{\mathbf{x}}} - \mathbf{m})' \, \tilde{s}_*^{-2} \, \boldsymbol{\delta}_*^{-2} (\tilde{\tilde{\mathbf{x}}} - \mathbf{m}) = \tilde{f}_2$$

aparține clasei $F(k, k(n-1))$.

În analogie cu cazul unidimensional, se pot considera statistici pentru compararea a două procese normale multidimensionale de generare de date. Astfel, dacă $\tilde{\mathbf{x}}_{11}, \ldots, \tilde{\mathbf{x}}_{1n_1}$ este o selecție de volum n_1 din populația vectorului aleator $\tilde{\mathbf{x}}_1 \in N(\mathbf{m}_1, \boldsymbol{\delta}_1^2)$, iar $\tilde{\mathbf{x}}_{21}, \ldots, \tilde{\mathbf{x}}_{2n_2}$ este o selecție de volum n_2 din populația vectorului aleator $\tilde{\mathbf{x}}_2 \in N(\mathbf{m}_2, \boldsymbol{\delta}_2^2)$, cele două selecții fiind independente, definim următoarele statistici:

$$\tilde{\tilde{\mathbf{x}}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \tilde{\mathbf{x}}_{ij}, \quad i = 1, 2, \tag{13.9.27}$$

$$\tilde{\mathbf{s}}_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (\tilde{\mathbf{x}}_{ij} - \tilde{\tilde{\mathbf{x}}}_i)(\tilde{\mathbf{x}}_{ij} - \tilde{\tilde{\mathbf{x}}}_i)', \quad i = 1, 2, \tag{13.9.28}$$

$$\tilde{\mathbf{s}}_i^2(\mathbf{m}_i) = \frac{1}{n_i} \sum_{j=1}^{n_i} (\tilde{\mathbf{x}}_{ij} - \mathbf{m}_i)(\tilde{\mathbf{x}}_{ij} - \mathbf{m}_i)', \quad i = 1, 2, \tag{13.9.29}$$

$$\tilde{s}_{*i}^2 = \frac{1}{k(n_i - 1)} \sum_{j=1}^{n_i} (\tilde{\mathbf{x}}_{ij} - \tilde{\tilde{\mathbf{x}}}_i)' \boldsymbol{\delta}_{*i}^{-2} (\tilde{\mathbf{x}}_{ij} - \tilde{\tilde{\mathbf{x}}}_i), \quad i = 1, 2, \tag{13.9.30}$$

$$\tilde{s}_{*i}^2(\mathbf{m}_i) = \frac{1}{kn_i} \sum_{j=1}^{n_i} (\tilde{\mathbf{x}}_{ij} - \mathbf{m}_i)' \boldsymbol{\delta}_{*i}^{-2} (\tilde{\mathbf{x}}_{ij} - \mathbf{m}_i), \quad i = 1, 2, \tag{13.9.31}$$

$$\tilde{\mathbf{s}}^2 = \frac{(n_1 - 1)\tilde{\mathbf{s}}_1^2 + (n_2 - 1)\tilde{\mathbf{s}}_2^2}{n_1 + n_2 - 2}, \tag{13.9.32}$$

$$\widetilde{s}_*^2 = \frac{(n_1 - 1)\,\widetilde{s}_{*1}^2 + (n_2 - 1)\widetilde{s}_{*2}^2}{n_1 + n_2 - 2}, \tag{13.9.33}$$

$$\widetilde{s}^2(m_1,\ m_2) = \frac{n_1 \widetilde{s}_1^2(m_1) + n_2 \widetilde{s}_2^2(m_2)}{n_1 + n_2}, \tag{13.9.34}$$

și

$$\widetilde{s}_*^2(m_1, m_2) = \frac{n_1 \widetilde{s}_{*1}^2(m_1) + n_2 \widetilde{s}_{*2}^2(m_2)}{n_1 + n_2}, \tag{13.9.35}$$

unde

$$\delta_{*i}^2 = \frac{1}{\sigma_{*i}^2}\,\delta_i^2, \quad \sigma_{*i}^2 = (\det \delta_i^2)^{1/k}, \quad i = 1, 2, \tag{13.9.36}$$

iar δ_{*i}^{-2} este inversa matricei δ_{*i}^2, $i = 1, 2$.

Teorema 13.9.5. *Cu notațiile și în condițiile specificate mai sus avem:*

(a) Statistica

$$\widetilde{q}_1 = (\widetilde{\bar{x}}_1 - \widetilde{\bar{x}}_2 - m_1 + m_2)' \left(\frac{1}{n_1}\,\delta_1^2 + \frac{1}{n_2}\,\delta_2^2\right)^{-1} (\widetilde{\bar{x}}_1 - \widetilde{\bar{x}}_2 - m_1 + m_2) \tag{13.9.37}$$

aparține clasei $\chi^2(k)$;

(b) Dacă, în plus, $\delta_1^2 = \delta_2^2 = \delta^2$, *atunci statistica*

$$\widetilde{f}_1 = \frac{n_1 n_2 (n_1 + n_2 - k - 1)}{k(n_1 + n_2)(n_1 + n_2 - 2)} (\widetilde{\bar{x}}_1 - \widetilde{\bar{x}}_2 - m_1 + m_2)' \widetilde{s}^{-2} (\widetilde{\bar{x}}_1 - \widetilde{\bar{x}}_2 - m_1 + m_2) \tag{13.9.38}$$

aparține clasei $F(k, n_1 + n_2 - k - 1)$;

(c) Dacă, în plus, $\sigma_{*1}^2 = \sigma_{*2}^2 = \sigma_*^2$, *atunci statistica*

$$\widetilde{q}_2 = \frac{k(n_1 + n_2 - 2)}{\sigma_*^2}\,\widetilde{s}_*^2 \tag{13.9.39}$$

aparține clasei $\chi^2(k(n_1+n_2-2))$, *statistica* $\tilde{s}_*^2$ *aparține clasei* $\gamma 2(1/\sigma_*^2,\ k(n_1+n_2-2))$, *statistica*

$$\tilde{q}_3 = \frac{k(n_1+n_2)}{\sigma_*^2}\,\tilde{s}_*^2(m_1,\ m_2) \tag{13.9.40}$$

aparține clasei $\chi^2(k(n_1+n_2))$, *statistica* $\tilde{s}_*^2(m_1, m_2)$ *aparține clasei* $\gamma 2(1/\sigma_*^2,\ k(n_1+n_2))$, *iar statistica*

$$\tilde{t}_2 = \frac{1}{k}\,(\tilde{\bar{x}}_1-\tilde{\bar{x}}_2-m_1+m_2)'\left[\tilde{s}_*^2\left(\frac{1}{n_1}\,\delta_{*1}^2+\frac{1}{n_2}\,\delta_{*2}^2\right)^{-1}(\tilde{\bar{x}}_1-\tilde{\bar{x}}_2-\right.$$

$$\left.-m_1+m_2)\right] \tag{13.9.41}$$

aparține clasei $F(k,\ k(n_1+n_2-2))$;

(d) Dacă, în plus, $\delta_1^2=\delta_2^2=\delta^2$, *atunci statistica*

$$\tilde{w}_1 = (n_1+n_2-2)\tilde{s}^2 \tag{13.9.42}$$

aparține clasei $W(\delta^{-2},\ n_1+n_2-2)$, *statistica* $\tilde{s}^2$ *aparține clasei* $W((n_1+n_2-2)\delta^{-2},\ n_1+n_2-2)$, *statistica*

$$\tilde{w}_2 = (n_1+n_2)\,\tilde{s}^2(m_1, m_2) \tag{13.9.43}$$

aparține clasei $W((n_1+n_2)\,\delta^{-2},\ n_1+n_2)$, *iar statistica* $\tilde{s}^2(m_1, m_2)$ *aparține clasei* $W((n_1+n_2)\,\delta^{-2},\ n_1+n_2)$.

Demonstrație. (a) Din teorema 15.9.3 rezultă că $\tilde{\bar{x}}_i$ aparține clasei $N\left(m_i,\ \frac{1}{n_i}\,\delta_i^2\right)$ pentru $i=1,\ 2$. Deoarece $\tilde{x}_{11},\ \ldots,\ \tilde{x}_{1n_1}$, $\tilde{x}_{21},\ \ldots,\ \tilde{x}_{2n_2}$ sînt vectori aleatori independenți, rezultă că $\tilde{\bar{x}}_1$ și $\tilde{\bar{x}}_2$ sînt de asemenea independenți și deci $\tilde{\bar{x}}_1-\tilde{\bar{x}}_2$ aparține clasei $N\left(m_1-m_2,\ \frac{1}{n_1}\,\delta_1^2+\frac{1}{n_2}\,\delta_2^2\right)$. Afirmația se deduce acum imediat din teorema 13.2.3.

Scanned with OKEN Scanner

(b) Din demonstrația punctului *(a)* urmează în cazul nostru că $\tilde{\bar{\mathbf{x}}}_1 - \tilde{\bar{\mathbf{x}}}_2$ aparține clasei $N\left(\mathbf{m}_1 - \mathbf{m}_2, \frac{n_1 + n_2}{n_1 n_2}\delta^2\right)$ și deci vectorul aleator $\tilde{\mathbf{y}}$ definit de

$$\tilde{\mathbf{y}} = \sqrt{\frac{n_1 + n_2}{n_1 n_2}}\,(\tilde{\bar{\mathbf{x}}}_1 - \tilde{\bar{\mathbf{x}}}_2 - \mathbf{m}_1 + \mathbf{m}_2)$$

aparține clasei $N(\mathbf{0},\ \delta^2)$.

Pe de altă parte, din teorema 13.9.3, rezultă că statistica $(n_i - 1)\tilde{s}_i^2$ aparține clasei $W(\delta^{-2},\ n_i - 1)$ pentru $i = 1, 2$. Cum $\tilde{s}_1^2$ și $\tilde{s}_2^2$ sînt independente, rezultă că suma

$$(n_1 - 1)\tilde{s}_1^2 + (n_2 - 1)\tilde{s}_2^2 = (n_1 + n_2 - 2)\tilde{s}^2$$

aparține clasei $W(\delta^{-2},\ n_1 + n_2 - 2)$.

Dacă observăm în plus că $\tilde{\mathbf{x}}_i$ și $\tilde{s}_i^2$ sînt independente pentru fiecare $i = 1, 2$, se deduce ușor că $\tilde{\mathbf{y}}$ și $\tilde{s}^2$ sînt independente și deci, conform teoremei 13.7.2, variabila aleatoare

$$\frac{n_1 + n_2 - k - 1}{k(n_1 + n_2 - 2)}\,\tilde{\mathbf{y}}'\tilde{s}^{-2}\tilde{\mathbf{y}} = \tilde{f}_1$$

aparține clasei $F(k, n_1 + n_2 - k - 1)$.

(c) Conform teoremei 13.9.3, statistica $k(n_i - 1)\tilde{s}_{*i}^2/\sigma_*^2$ aparține clasei $\chi^2(k(n_i - 1))$ pentru $i = 1, 2$. Deoarece $\tilde{s}_{*1}^2$ și $\tilde{s}_{*2}^2$ sînt independente, rezultă că suma

$$\frac{k(n_1 - 1)\tilde{s}_{*1}^2 + k(n_2 - 1)\tilde{s}_{*2}^2}{\sigma_*^2} = \frac{k(n_1 + n_2 - 2)}{\sigma_*^2}\,\tilde{s}_*^2 = \tilde{q}_2$$

aparține clasei $\chi^2(k(n_1 + n_2 - 2))$, de unde rezultă imediat că $\tilde{s}_*^2$ aparține clasei $\gamma 2(1/\sigma_*^2,\ k(n_1 + n_2 - 2))$.

Apoi, utilizînd din nou teorema 13.9.3 rezultă că statistica $kn_i\tilde{s}_{*i}^2(\mathbf{m}_i)/\sigma_*^2$ aparține clasei $\chi^2(kn_i)$ pentru $i = 1, 2$. Deoarece $\tilde{s}_{*1}^2(\mathbf{m}_1)$ și $\tilde{s}_{*2}^2(\mathbf{m}_2)$ sînt independente, urmează că suma

$$\frac{kn_1\tilde{s}_{*1}^2(\mathbf{m}_1) + kn_2\tilde{s}_{*2}^2(\mathbf{m}_2)}{\sigma_*^2} = \frac{k(n_1 + n_2)}{\sigma_*^2}\,\tilde{s}_*^2(\mathbf{m}_1, \mathbf{m}_2) = \tilde{q}_3$$

aparține clasei $\chi^2(k(n_1 + n_2))$ și deci $\tilde{s}_*^2(\mathbf{m}_1, \mathbf{m}_2)$ aparține clasei $\gamma 2(1/\sigma_*^2,\ k(n_1 + n_2))$.

Scanned with OKEN Scanner

Pentru a demonstra și ultima afirmație, să observăm că statistica

$$\tilde{q}_1 = \frac{1}{\sigma_*^2}(\tilde{\bar{x}}_1 - \tilde{\bar{x}}_2 - m_1 + m_2)'\left(\frac{1}{n_1}\delta_{*1}^2 + \frac{1}{n_2}\delta_{*2}^2\right)^{-1}(\tilde{\bar{x}}_1 - \tilde{\bar{x}}_2 - m_1 + m_2)$$

aparține clasei $\chi^2(k)$, conform punctului *(a)* al teoremei noastre. Deoarece statisticile $\tilde{q}_1$ și $\tilde{q}_2$ sînt independente, urmează că raportul

$$\frac{\tilde{q}_1/k}{\tilde{q}_2/k(n_1+n_2-2)} = \tilde{f}_2$$

aparține clasei $F(k, k(n_1+n_2-2))$.

(d) Să observăm mai întîi că statistica $(n_i-1)\tilde{s}_i^2$ aparține clasei $W(\delta^{-2}, n_i-1)$ pentru $i = 1, 2$, conform teoremei 13.9.3. Cum $\tilde{s}_1^2$ și $\tilde{s}_2^2$ sînt independente, urmează că suma

$$(n_1-1)\tilde{s}_1^2 + (n_2-1)\tilde{s}_2^2 = (n_1+n_2-2)\tilde{s}^2 = \tilde{w}_1$$

aparține clasei $W(\delta^{-2}, n_1+n_2-2)$ și deci $\tilde{s}^2$ aparține clasei $W((n_1+n_2-2)\,\delta^{-2}, n_1+n_2-2)$.

Analog, utilizînd din nou teorema 13.9.3, rezultă că $n_i\tilde{s}_i^2(m_i)$ aparține clasei $W(\delta^{-2}, n_i)$ pentru $i = 1, 2$. Deoarece $\tilde{s}_1^2(m_1)$ și $\tilde{s}_2^2(m_2)$ sînt independente, rezultă că suma

$$n_1\tilde{s}_1^2(m_1) + n_2\tilde{s}_2^2(m_2) = (n_1+n_2)\tilde{s}^2(m_1, m_2) = \tilde{w}_2$$

aparține clasei $W(\delta^{-2}, n_1+n_2)$ și deci $\tilde{s}^2(m_1, m_2)$ aparține clasei $W(n_1+n_2)\delta^{-2}, n_1+n_2)$.

Scanned with OKEN Scanner

Capitolul 14

FAMILII DE REPARTIȚII CONJUGATE

14.1. INTRODUCERE

Atît în problemele de inferență bayesiană, cît și în cele de decizie sîntem interesați ca repartițiile apriori și cele aposteriori să aparțină unei aceleiași familii de repartiții, situație care facilitează mult determinarea repartițiilor aposteriori, deoarece trebuie schimbate numai valorile parametrilor familiei de repartiții.

Vom spune că o familie de repartiții este *conjugată* cu un proces de generare de date de parametru $\boldsymbol{\theta}$, dacă orice repartiție din această familie, utilizată ca repartiție apriori pentru $\boldsymbol{\theta}$, produce repartiții aposteriori pentru $\boldsymbol{\theta}$ din aceeași familie.

În cele ce urmează nu vom prezenta o teorie generală a conjugării *, mulțumindu-ne să indicăm familiile de repartiții conjugate cu procesele de generare de date uzuale. În acest scop vom considera o familie de repartiții conjugate cu un anumit proces de generare de date de parametru $\boldsymbol{\theta}$ și vom determina repartiția aposteriori corespunzătoare. Pentru aceasta vom scrie funcția de verosimilitate $L(\boldsymbol{\theta}/n, \mathbf{x}_1, \ldots, \mathbf{x}_n)$ a procesului considerat, vom considera o statistică suficientă $\tilde{\mathbf{x}}_T$ pentru $\boldsymbol{\theta}$ și vom scrie repartiția elementară pentru $(\tilde{\boldsymbol{\theta}}, \tilde{\mathbf{x}}_T)$ sub forma

$$f(\boldsymbol{\theta}, \mathbf{x}_T) = f(\boldsymbol{\theta})\, f(\mathbf{x}_T \mid \boldsymbol{\theta}) = q(\mathbf{x}_1, \ldots, \mathbf{x}_n)\, f(\boldsymbol{\theta})\, L(\boldsymbol{\theta} \mid \mathbf{x}_T) \qquad (14.1.1)$$

pentru a *recunoaște* pe acest produs partea care depinde efectiv de parametru, adică $f(\boldsymbol{\theta})\, L(\boldsymbol{\theta}/\mathbf{x}_T)$.

Să observăm de asemenea că avem

$$f(\boldsymbol{\theta} \mid \mathbf{x}_T) = \frac{f(\boldsymbol{\theta}, \mathbf{x}_T)}{f(\mathbf{x}_T)} = \frac{q(\mathbf{x}_1, \ldots, \mathbf{x}_n)\, f(\boldsymbol{\theta})\, L(\boldsymbol{\theta} \mid \mathbf{x}_T)}{f(\mathbf{x}_T)} \qquad (14.1.2)$$

care conține factorul $q(\mathbf{x}_1, \ldots, \mathbf{x}_n)/f(\mathbf{x}_T)$ *independent* de $\boldsymbol{\theta}$. Din relațiile (14.1.1) și (14.1.2) rezultă că atît $f(\boldsymbol{\theta}, \mathbf{x}_T)$ cît și $f(\boldsymbol{\theta}/\mathbf{x}_T)$

---

* Cititorul interesat poate urmări o astfel de teorie în Raiffa, H., Schlaifer, R. (1961).

sînt proporționali cu $f(\theta)\, L(\theta/\mathbf{x}_T)$, ceea ce ne permite să *recunoaștem* repartiția elementară aposteriori $f(\theta/\mathbf{x}_T)$ din (14.1.1), dacă ținem seama că $f(\theta/\mathbf{x}_T)$ *trebuie* să aparțină aceleiași familii de repartiții elementare ca și $f(\theta)$.

Metoda de determinare a familiilor de repartiții conjugate pe care am expus-o mai sus este numită *metoda recunoașterii* și va fi utilizată aproape fără excepție în cele ce urmează.

14.2. PROCESUL BERNOULLI

Să considerăm un proces Bernoulli de generare de date de parametru $p \in (0, 1)$. Dacă procesul de oprire este neinformativ pentru p, atunci funcția de verosimilitate este de forma

$$L(p \mid n, x_1, \ldots, x_n) = \begin{cases} p^r (1-p)^{n-r} & \text{pentru } r \in \{0, 1, \ldots, n\}, \\ 0 & \text{în caz contrar,} \end{cases} \tag{14.2.1}$$

unde $x_1, \ldots, x_n \in \{0, 1\}$ sînt rezultatele celor n observații, iar $r = x_1 + \ldots + x_n$.

Vom arăta în cele ce urmează că o familie de repartiții elementare conjugată cu acest proces este formată de densitățile de probabilitate $f_\beta(p \mid \rho, \nu)$, corespunzătoare repartiției beta de parametrii ρ și ν.

Într-adevăr, repartiția elementară pentru vectorul aleator $(\tilde{p}, \tilde{r}, \tilde{n})$ are valori nenule de forma

$$f(p, r, n) = f(p)\, f(r, n, p),$$

adică

$$f(p, r, n) = L(p \mid r, n)\, f_\beta(p \mid \rho, \nu) =$$

$$= \frac{1}{B(\rho, \nu)}\, p^{\rho + r - 1} (1-p)^{\nu + n - r - 1}. \tag{14.2.2}$$

Deoarece partea care depinde de statistica suficientă $(\tilde{n}, \tilde{r})$ pentru p este $p^{\rho + r - 1} (1-p)^{\nu + n - r - 1}$, rezultă că densitatea de probabilitate aposteriori pentru $\tilde{p}$ trebuie să fie $f_\beta(p \mid \rho + r,$

$\nu + n - r$), conform metodei recunoașterii expusă în § 14.1 (vezi formula (14.1.2)).

Din cele de mai sus rezultă că „actualizarea" densității de probabilitate apriori $f_\beta(p \mid \rho, \nu)$ se reduce la înlocuirea parametrilor inițiali ρ, ν cu valorile $\rho + r$ și $\nu + n - r$.

După cum este cunoscut, dacă n este fixat, atunci funcția de masă a lui $\tilde{r}$ este $f_b(r \mid n, p)$. Dacă densitatea de probabilitate apriori pentru $\tilde{p}$ este $f_\beta(p \mid \rho, \nu)$, atunci rezultă că repartiția elementară a vectorului aleator $(\tilde{r}, \tilde{p})$ are valori nenule de forma

$$f(r, p) = f_b(r \mid n, p)\, f_\beta(p \mid \rho, \nu) =$$

$$= \frac{n!}{r!(n-r)!} p^r (1-p)^{n-r} \cdot \frac{p^{\rho-1}(1-p)^{\nu-1}}{B(\rho, \nu)} =$$

$$= \frac{n!}{r!(n-r)!} \cdot \frac{\Gamma(\rho+\nu)}{\Gamma(\rho)\ \Gamma(\nu)} p^{\rho+r-1}(1-p)^{n-r+\nu-1} =$$

$$= \frac{\Gamma(n+1)\ \Gamma(\rho+\nu)}{\Gamma(r+1)\ \Gamma(n-r+1)\ \Gamma(\rho)\ \Gamma(\nu)} p^{\rho+r-1}(1-p)^{n-r+\nu-1}$$

și deci funcția de masă (marginală) pentru $\tilde{r}$ are valori nenule de forma

$$f(r) = \frac{\Gamma(n+1)\ \Gamma(\rho+\nu)\ \Gamma(\rho+r)\ \Gamma(n-r+\nu)}{\Gamma(r+1)\ \Gamma(n-r+1)\ \Gamma(\rho)\ \Gamma(\nu)\ \Gamma(\rho+n+\nu)}, \tag{14.2.3}$$

unde $r \in \{0, 1, \ldots, n\}$.

Funcția de masă (14.2.3) pune în evidență o repartiție discretă numită *repartiția beta-binomială de parametrii* ρ, ν *și* n. Funcția de masă (14.2.3) corespunzătoare acestei repartiții este notată prin $f_{\beta b}(r \mid \rho, \nu, n)$, iar mulțimea variabilelor aleatoare cu această repartiție prin $\beta b(\rho, \nu, n)$.

Se poate vedea ușor că dacă $\tilde{x} \in \beta b(\rho, \nu, n)$, atunci

$$M(\tilde{x}) = \frac{n\rho}{\nu + \rho}, \qquad V(\tilde{x}) = \frac{n\rho\nu(n+\rho+\nu)}{(\rho+\nu)^2\,(\rho+\nu+1)}. \tag{14.2.4}$$

Într-adevăr, ținînd seama de proprietățile valorii medii condiționate, rezultă

$$M_{\beta b}(\tilde{r}\,|\,\rho, \nu, n) = M_{\beta}(M_b(\tilde{r}\,|\,n,\ \tilde{p})\,|\,\rho, \nu) = M_{\beta}(n\tilde{p}\,|\,\rho, \nu) =$$

$$= n M_{\beta}(\tilde{p}\,|\,\rho, \nu) = \frac{n\rho}{\rho + \nu},$$

dacă se utilizează rezultatele cunoscute (12.3.8) și (13.3.5).

Pentru a demonstra cea de a doua relație din (14.2.4), să observăm mai întîi că putem scrie

$$V_{\beta b}(\tilde{r}\,|\,\rho, \nu, n) = M_{\beta}(V_b(\tilde{r}\,|n, \tilde{p})\ |\ \rho, \nu) + V_{\beta}(M_b(\tilde{r}\,|n, \tilde{p})\,|\,\rho, \nu),$$

conform unei proprietăți cunoscute a dispersiei (vezi formula (6.5.6)). Deoarece

$$M_{\beta}(V_b(\tilde{r}\,|n, \tilde{p})\,|\,\rho, \nu) = M_{\beta}(n\tilde{p}(1 - \tilde{p})\,|\,\rho, \nu) =$$

$$= n M_{\beta}(\tilde{p}\,|\,\rho, \nu) - n M_{\beta}(\tilde{p}^2\,|\,\rho, \nu) = \frac{n\rho}{\rho + \nu} - \frac{n\rho(\rho + 1)}{(\rho + \nu)(\rho + \nu + 1)}$$

și

$$V_{\beta}(M_b(\tilde{r}\,|n, \tilde{p})\,|\,\rho, \nu) = V_{\beta}(n\tilde{p}\,|\,\rho, \nu) = \frac{n^2\rho\nu}{(\rho + \nu)^2(\rho + \nu + 1)},$$

rezultă

$$V_{\beta b}(\tilde{r}\,|\,\rho, \nu, n) = \frac{n\rho\nu(n + \rho + \nu)}{(\rho + \nu)^2(\rho + \nu + 1)},$$

ceea ce demonstrează complet relațiile (14.2.4).

După cum este cunoscut, dacă r este fixat, atunci funcția de masă a lui $\tilde{n}$ este $f_{Pa}(n \mid r, p)$. Dacă se consideră din nou pentru $\tilde{p}$ densitatea de probabilitate $f_{\beta}(p\,|\,\rho, \nu)$, rezultă că repartiția elementară a vectorului aleator $(\tilde{n}, \tilde{p})$ are valori nenule de forma

$$f(n, p) = f_{\beta}(p\,|\,\rho, \nu) f_{Pa}(n\,|\,r, p) =$$

$$= \frac{p^{\rho-1}(1 - p)^{\nu-1}}{B(\rho, \nu)} \cdot \frac{(n - 1)!}{(r - 1)!\,(n - r)!} p^r (1 - p)^{n-r} =$$

$$= \frac{\Gamma(\rho+\nu)}{\Gamma(\rho)\,\Gamma(\nu)} \cdot \frac{\Gamma(n)}{\Gamma(r)\,\Gamma(n-r+1)}\, p^{\rho+r-1}(1-p)^{n-r+\nu-1}$$

pentru $n \in \{r, r+1, \ldots\}$ și $p \in (0, 1)$. Rezultă deci că funcția de masă (marginală) pentru $\tilde{n}$ are valori nenule de forma

$$f(n) = \frac{\Gamma(\rho+\nu)\,\Gamma(n)\,\Gamma(\rho+r)\,\Gamma(n-r+\nu)}{\Gamma(\rho)\,\Gamma(\nu)\,\Gamma(r)\,\Gamma(n-r+1)\,\Gamma(\rho+\nu+n)} \tag{14.2.5}$$

pentru $n \in \{r+1, r+2, \ldots\}$. Funcția de masă (14.2.5) definește o repartiție discretă numită *repartiția beta-Pascal de parametrii r, ρ și ν*. Funcția de masă (14.2.5) a unei repartiții beta-Pascal va fi notată prin $f_{\beta Pa}(n|r, \rho, \nu)$, iar mulțimea variabilelor aleatoare discrete cu această repartiție prin $\beta Pa(r, \rho, \nu)$.

Se poate vedea ușor că dacă $\tilde{x} \in \beta Pa(r, \rho, \nu)$, atunci

$$M(\tilde{x}) = \frac{r(\rho+\nu-1)}{\rho-1}, \tag{14.2.6}$$

cu condiția ca $\rho > 1$ și

$$V(\tilde{x}) = \frac{r\,\nu(r+\rho-1)(\nu+\rho-1)}{(\rho-1)^2(\rho-2)}, \tag{14.2.7}$$

cu condiția ca $\rho > 2$.

Într-adevăr, pentru a demonstra (14.2.6), vom utiliza din nou proprietatea valorii medii condiționate folosită mai sus și vom obține:

$$M_{\beta Pa}(\tilde{n}|r, \rho, \nu) = M_{\beta}(M_{Pa}(\tilde{n}|r, \tilde{p})|\rho, \nu) = M_{\beta}\left(\frac{r}{\tilde{p}}\,\middle|\,\rho, \nu\right) =$$

$$= r\,M_{\beta}\left(\frac{1}{\tilde{p}}\,\middle|\,\rho, \nu\right) = \frac{r}{B(\rho, \nu)}\int_0^1 p^{\rho-2}(1-p)^{\nu-1}\,\mathrm{d}p =$$

$$= r\,\frac{B(\rho-1, \nu)}{B(\rho, \nu)} = \frac{r(\rho+\nu-1)}{\rho-1}$$

cu condiția ca $\rho > 1$, ceea ce demonstrează (14.2.6).

Scanned with OKEN Scanner

Pentru a arăta (14.2.7), utilizăm din nou proprietatea (6.5.6) a dispersiei și obținem:

$$V_{\beta Pa}(\tilde{n}|r,\ \rho,\ \nu) = M_\beta(V_{Pa}(\tilde{n}|r,\tilde{p})|\rho,\ \nu) + V_\beta(M_{Pa}(\tilde{n}|r,\tilde{p})|\rho,\ \nu).$$

Deoarece, pentru $\rho > 2$, avem

$$M_\beta(V_{Pa}(\tilde{n}|r,\tilde{p})|\rho,\ \nu) = M_\beta\left(\frac{r(1-\tilde{p})}{\tilde{p}^2}\middle|\rho,\ \nu\right) =$$

$$= \frac{r}{B(\rho,\ \nu)}\int_0^1 p^{\rho-3}(1-p)^\nu\,dp = \frac{r\,B(\rho-2,\ \nu+1)}{B(\rho,\ \nu)} =$$

$$= \frac{r\,\nu(\rho+\nu-1)}{(\rho-1)\,(\rho-2)}$$

și

$$V_\beta(M_{Pa}(\tilde{n}|r,\tilde{p})|\rho,\ \nu) = V_\beta\left(\frac{r}{\tilde{p}}\middle|\ \rho,\ \nu\right) = M_\beta\left(\frac{r^2}{\tilde{p}^2}\middle|\ \rho,\ \nu\right) -$$

$$- M_\beta^2\left(\frac{r}{\tilde{p}}\middle|\ \rho,\ \nu\right) = \frac{r^2}{B(\rho,\ \nu)}\int_0^1 p^{\rho-3}(1-p)^{\nu-1}\,dp -$$

$$- \frac{r^2(\rho+\nu-1)^2}{(\rho-1)^2} = \frac{r^2\nu(\rho+\nu-1)}{(\rho-1)^2\,(\rho-2)},$$

rezultă că obținem

$$V_{\beta Pa}(\tilde{n}|r,\ \rho,\ \nu) = \frac{r\,\nu(r+\rho-1)\,(\nu+\rho-1)}{(\rho-1)^2\,(\rho-2)},$$

ceea ce demonstrează și formula (14.2.7).

Scanned with OKEN Scanner

14.3. PROCESUL BERNOULLI MULTIDIMENSIONAL

Să considerăm un proces Bernoulli k—dimensional $(\widetilde{\mathbf{x}}_i)$ de generare de date de parametru $\mathbf{p} = (p_1, \ldots, p_k)'$. Funcția sa de verosimilitate are valori nenule de forma

$$L(\mathbf{p} \mid n, \mathbf{x}_1, \ldots, \mathbf{x}_n) = p_0^{r_0} p_1^{r_1} \ldots p_k^{r_k} \tag{14.3.1}$$

pentru toți $r_j \in \{0, 1, \ldots, n\}$ pentru care

$$r_0 + r_1 + \ldots + r_k = n;\ p_0 = 1 - (p_1 + \ldots p_k), \tag{14.3.2}$$

unde

$$r_j = \sum_{i=1}^{n} x_{ji}, \tag{14.3.3}$$

iar x_{ji} este componenta de rang j, $1 \leqslant j \leqslant k$, a vectorului $\mathbf{x}_i$, $1 \leqslant i \leqslant n$. Vectorul cu componentele r_j, $1 \leqslant j \leqslant k$, va fi notat prin $\mathbf{r}$ și deci putem scrie (14.3.3) și sub forma

$$\mathbf{r} = \sum_{i=1}^{n} \mathbf{x}_i. \tag{14.3.4}$$

Repartiția conjugată cu procesul Bernoulli multidimensional este repartiția beta de dimensiune corespunzătoare procesului. Mai precis, densitățile de probabilitate $f_\beta(\mathbf{p} \mid \boldsymbol{\rho}, \nu)$ definite în § 13.5 constituie o familie conjugată cu procesul Bernoulli de generare de date de parametru $\mathbf{p}$. Într-adevăr, dacă adoptăm pentru vectorul aleator $\widetilde{\mathbf{p}} = (\widetilde{p}_1, \ldots, \widetilde{p}_k)'$ densitatea de probabilitate apriori $f_\beta(\mathbf{p} \mid \boldsymbol{\rho}, \nu)$, atunci rezultă că repartiția elementară a vectorului aleator $(\widetilde{n}, \widetilde{\mathbf{r}}, \widetilde{\mathbf{p}})$ este de forma

$$f(n, \mathbf{r}, \mathbf{p}) = f_\beta(\mathbf{p} \mid \boldsymbol{\rho}, \nu)\, L(\mathbf{p} \mid n, \mathbf{r}).$$

Valorile nenule ale acestei repartiții elementare sînt deci de forma

$$\frac{1}{B_k(\rho, \nu)} p_1^{\rho_1 + r_1 - 1} \ldots p_k^{\rho_k + r_k - 1} (1 - p_1 - \ldots - p_k)^{\nu + r_0 - 1}, \tag{14.3.5}$$

ceea ce arată că densitatea de probabilitate aposteriori pentru $\widetilde{\mathbf{p}}$ *trebuie* să fie $f_\beta(\mathbf{p} \mid \boldsymbol{\rho} + \mathbf{r},\ \nu + r_0)$, unde

$$r_0 = n - (r_1 + \ldots + r_k). \tag{14.3.6}$$

Scanned with OKEN Scanner

Dacă se fixează n, atunci este cunoscut că funcția de masă a lui $\tilde{\mathbf{r}}$ este $f_{mu}(\mathbf{r} \mid n, \mathbf{p})$. Dacă adoptăm deci pentru $\tilde{\mathbf{p}}$ densitatea de probabilitate apriori $f_{\beta}(\mathbf{p} \mid \boldsymbol{\rho}, \nu)$, atunci rezultă că funcția de masă a lui $\tilde{\mathbf{r}}$ va avea valori nenule de forma

$$f(\mathbf{r}) = \frac{f(\mathbf{r}, \mathbf{p})}{f(\mathbf{p} \mid \mathbf{r})} = f_{mu}(\mathbf{r} \mid n, \mathbf{p}) \cdot \frac{f_{\beta}(\mathbf{p} \mid \boldsymbol{\rho}, \nu)}{f_{\beta}(\mathbf{p} \mid \boldsymbol{\rho} + \mathbf{r}, \nu + r_0)} =$$

$$= \frac{n!}{r_0!\, r_1! \ldots r_k!} p_0^{r_0} p_1^{r_1} \ldots p_k^{r_k} \cdot \frac{B_k(\boldsymbol{\rho} + \mathbf{r}, \nu + r_0)}{B_k(\boldsymbol{\rho}, \nu)} p_0^{-r_0} \ldots p_k^{-r_k}$$

sau

$$f(\mathbf{r}) = \frac{B_k(\boldsymbol{\rho} + r, \nu + r_0)}{B_k(\boldsymbol{\rho}, \nu)} \cdot \frac{\Gamma(n+1)}{\Gamma(r_0+1) \ldots \Gamma(r_k+1)} \tag{14.3.7}$$

pentru orice $r_j \geqslant 0$, $0 \leqslant j \leqslant k$, care verifică relația (14.3.6).

Repartiția definită de funcția de masă (14.3.7) va fi numită *repartiția beta-multinomială de parametrii* $\boldsymbol{\rho}, \nu$ *și* n, iar funcția de masă (14.3.7) va fi notată prin $f_{\beta mu}(\mathbf{r} \mid \boldsymbol{\rho}, \nu, n)$. Mulțimea variabilelor aleatoare cu această repartiție va fi notată prin $\beta\, mu(\boldsymbol{\rho}, \nu, n)$.

Se poate arăta ușor că dacă $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_k)'$ este un vector aleator din clasa $\beta mu(\boldsymbol{\rho}, \nu, n)$, atunci

$$M(\tilde{x}_j) = \frac{n \rho_j}{\rho_1 + \ldots + \rho_k + \nu}, \tag{14.3.8}$$

$$V(\tilde{x}_j) = \frac{n \rho_j (n + \rho_1 + \ldots + \rho_k + \nu)}{(\nu + \rho_1 + \ldots + \rho_k)^2} \cdot$$

$$\cdot \frac{(\nu + \rho_1 + \ldots + \rho_{j-1} + \rho_{j+1} + \ldots + \rho_k)}{(\nu + \rho_1 + \ldots + \rho_k + 1)}, \tag{14.3.9}$$

și

$$\operatorname{cov}(\tilde{x}_i, \tilde{x}_j) = \frac{-n\, \rho_i \rho_j (n + \nu + \rho_1 + \ldots + \rho_k)}{(\rho_1 + \ldots + \rho_k + \nu)^2 (\rho_1 + \ldots + \rho_k + \nu + 1)}, \quad i \neq j. \tag{14.3.10}$$

Într-adevăr, pentru a arăta (14.3.8), să observăm că

$$M_{\beta mu}(\tilde{r}_j|\,\rho,\nu,n)=M_\beta(M_{mu}(\tilde{r}_j|n,\tilde{\mathbf{p}})|\rho,\nu)=M_\beta(n\tilde{p}_j|\,\rho,\nu)=$$

$$=n\,M_\beta(\tilde{p}_j|\,\rho,\nu)=\frac{n\,\rho_j}{\rho_1+\ldots+\rho_k+\nu},$$

conform cu rezultatele anterioare (12.4.13) și (13.5.8).

Pentru a arăta (14.3.9) putem utiliza din nou proprietatea (6.5.6) a dispersiei și obținem

$$V_{\beta mu}(\tilde{r}_j|\,\rho,\nu,n)=M_\beta(V_{mu}(\tilde{r}_j|\,n,\tilde{\mathbf{p}})\,|\,\rho,\nu)+$$

$$+V_\beta(M_{mu}(\tilde{r}_j|n,\tilde{\mathbf{p}})\,|\,\rho,\nu). \qquad (14.3.11)$$

Primul termen al sumei din (14.3.11) devine

$$M_\beta(V_{mu}(\tilde{r}_j|n,\tilde{\mathbf{p}})\,|\,\rho,\nu)=M_\beta(n\tilde{p}_j(1-\tilde{p}_j)\,|\,\rho,\nu)=$$

$$=n\,M_\beta(\tilde{p}_j|\,\rho,\nu)-n\,M_\beta(\tilde{p}_j^2|\,|\,\rho,\nu)= \qquad (14.3.12)$$

$$=\frac{n\,\rho_j}{\rho_1+\ldots+\rho_k+\nu}-\frac{n\,\rho_j(\rho_j+1)}{(\rho_1+\ldots+\rho_k+\nu)(\rho_1+\ldots+\rho_k+\nu+1)},$$

iar al doilea termen poate fi scris sub forma

$$V_\beta(M_{mu}(\tilde{r}_j|\,n,\tilde{\mathbf{p}})|\,\rho,\nu)=V_\beta(n\tilde{p}_j|\,\rho,\nu)=n^2V_\beta(\tilde{p}_j|\,\rho,\nu)=$$

$$=\frac{n^2\rho_j(\rho_1+\ldots+\rho_{j-1}+\rho_{j+1}+\ldots+\rho_k+\nu)}{(\rho_1+\ldots+\rho_k+\nu)^2(\rho_1+\ldots+\rho_k+\nu+1)}. \qquad (14.3.13)$$

Din relațiile (14.3.11), (14.3.12) și (14.3.13) rezultă imediat formula căutată (14.3.9).

Pentru demonstrarea formulei (14.3.10), vom utiliza proprietatea (6.5.9) a corelației. Obținem

$$\mathrm{cov}_{\beta mu}(\tilde{r}_i,\tilde{r}_j|\,\rho,\nu,n)=M_\beta(\mathrm{cov}_{mu}(\tilde{r}_i,\tilde{r}_j|\,n,\tilde{p})\,|\,\rho,\nu)+$$

$$+\mathrm{cov}_\beta(M_{mu}(\tilde{r}_i|\,n,\tilde{\mathbf{p}}),\,M_{mu}(\tilde{r}_j|\,n,\tilde{p})|\,\rho,\nu). \qquad (14.3.14)$$

Scanned with OKEN Scanner

Ținînd seama de rezultatele anterioare (12.4.14) și (13.5.11), primul termen al sumei din (14.3.14) devine

$$M_\beta(\mathrm{cov}_{mu}(\tilde{r}_i, \tilde{r}_j | n, \tilde{\mathbf{p}}) \mid \boldsymbol{\rho}, \nu) = M_\beta(-n\tilde{p}_i\tilde{p}_j | \boldsymbol{\rho}, \nu) =$$

$$= \frac{-n\,\rho_i\rho_j}{(\rho_1 + \ldots + \rho_k + \nu)(\rho_1 + \ldots + \rho_k + \nu + 1)}. \tag{14.3.15}$$

Analog, ținînd seama de rezultatele anterioare (12.4.13) și (13.5.12), al doilea termen al sumei din (14.3.14) devine

$$\mathrm{cov}_\beta(M_{mu}(\tilde{r}_i | n, \tilde{\mathbf{p}}), M_{mu}(\tilde{r}_j | n, \tilde{\mathbf{p}}) \mid \boldsymbol{\rho}, \nu) = \mathrm{cov}_\beta(n\tilde{p}_i, n\tilde{p}_j \mid \boldsymbol{\rho}, \nu) =$$

$$= \frac{-n^2\,\rho_i\,\rho_j}{(\rho_1 + \ldots + \rho_k + \nu)^2(\rho_1 + \ldots + \rho_k + \nu + 1)}. \tag{14.3.16}$$

Din relațiile (14.3.14), (14.3.15) și (14.3.16) urmează imediat formula căutată (14.3.10).

14.4. PROCESUL POISSON

Să considerăm un proces Poisson de generare de date $(\tilde{t}_i)$ de parametru $\lambda > 0$. Funcția de verosimilitate pentru acest proces are valori nenule de forma

$$L(\lambda \mid t, t_1, \ldots, t_r) = \lambda^r\, e^{-\lambda t} \tag{14.4.1}$$

pentru toate valorile $t_1, \ldots, t_r$ pozitive pentru care este îndeplinită condiția

$$t_1 + \ldots + t_r \leqslant t. \tag{14.4.2}$$

Este clar că dacă procesul de oprire este neinformativ pentru λ, rezultă că statistica $(\tilde{r}, \tilde{\tau})$ este suficientă pentru parametrul λ al procesului. Prin urmare, funcția de verosimilitate poate fi scrisă sub forma

$$L(\lambda \mid r, t) = \lambda^r\, e^{-\lambda t} \tag{14.4.3}$$

pentru oriee $t > 0$, avînd valori nule pentru $t \leqslant 0$.

Se poate arăta ușor că o familie de repartiții conjugate cu procesul Poisson de generare de date de parametru λ este repartiția gama. Într-adevăr, utilizînd pentru $\tilde{\lambda}$ o densitate de probabilitate apriori $f_\gamma(\lambda|\rho, \tau)$, rezultă că densitatea de probabilitate aposteriori pentru $\tilde{\lambda}$ depinde de λ numai prin intermediul produsului

$$f_\gamma(\lambda|\rho, \tau)\, L(\lambda|\, r, \, t) = \frac{\tau\, e^{-\lambda\tau}\,(\lambda\tau)^{\rho-1}}{\Gamma(\rho)}\,\lambda^r\, e^{-\lambda t} =$$

$$= \frac{\tau^\rho}{\Gamma(\rho)}\, e^{-\lambda(t+\tau)}\,\lambda^{\rho+r-1},$$

ceea ce arată că densitatea de probabilitate aposteriori pentru $\tilde{\lambda}$ trebuie să fie $f_\gamma(\lambda|\rho+r, \tau+t)$, cu condiția ca procesul de oprire utilizat să fie neinformativ.

Dacă se fixează durata de timp t în care se fac observațiile, atunci $\tilde{r}$ aparține clasei $Po(\lambda t)$. Dacă utilizăm pentru $\tilde{\lambda}$ o densitate de probabilitate apriori $f_\gamma(\lambda|\rho, \tau)$, atunci rezultă că funcția de masă a lui $\tilde{r}$ este

$$f(r) = \frac{f(r, \lambda)}{f(\lambda|\, r)} = f_{Po}(r|\, \lambda t)\cdot f_\gamma(\lambda|\, \rho, \, \tau)\cdot \frac{1}{f_\gamma(\lambda|\rho + r, \, \tau + t)} =$$

$$= \frac{\Gamma(\rho + r)}{\Gamma(\rho)}\cdot\frac{e^{-\lambda t}\,(\lambda t)^r}{\Gamma(r+1)}\cdot\frac{\tau\, e^{-\tau\lambda}\,(\lambda\tau)^{\rho-1}}{(\tau + t)\, e^{-\lambda(\tau+t)}[\lambda(\tau + t)]^{\rho+r-1}} =$$

$$= \frac{\Gamma(\rho + r)}{\Gamma(\rho)\,\Gamma(r + 1)}\left(\frac{\tau}{\tau + t}\right)^\rho\left(1 - \frac{\tau}{\tau + t}\right)^r. \qquad (14.4.4)$$

Pentru ρ număr natural avem

$$\frac{\Gamma(\rho + r)}{\Gamma(\rho)\,\Gamma(r + 1)} = \frac{(\rho + r - 1)\,!}{(\rho - 1)\,!\,r\,!} = C^{\rho-1}_{\rho+r-1}$$

și deci funcția de masă $f(r)$ dată de (14.4.4) este, conform formulei (12.3.22), chiar $f_{bn}\left(r|\, \rho, \, \frac{\tau}{\tau+t}\right)$.

Repartiția binomială negativă (10.3.22) poate fi deci generalizată, luînd funcția sa de masă de forma

$$f_{bn}(r \mid \rho, p) = \begin{cases} \dfrac{\Gamma(\rho + r)}{\Gamma(\rho)\,\Gamma(r+1)}\, p^{\rho}(1-p)^{r} & \text{pentru } r \in \{0, 1, \ldots\}, \\ 0 & \text{în caz contrar,} \end{cases} \tag{14.4.5}$$

unde $\rho > 0$ nu este neapărat întreg.

Media lui $\tilde{r}$ se obține prin calcul direct în modul următor:

$$M_{bn}(\tilde{r} \mid \rho,\ p) = \sum_{r=0}^{\infty} r\, \frac{\Gamma(\rho + r)}{\Gamma(\rho)\,\Gamma(r+1)}\, p^{\rho}(1-p)^{r} =$$

$$= \sum_{r=1}^{\infty} \frac{\Gamma(\rho + r)}{\Gamma(\rho)\,\Gamma(r)}\, p^{\rho}\,(1-p)^{r} =$$

$$= \sum_{r=0}^{\infty} \frac{\Gamma(\rho + r + 1)}{\Gamma(\rho)\,\Gamma(r+1)}\, p^{\rho}\,(1-p)^{r+1} =$$

$$= (1-p) \sum_{r=0}^{\infty} \frac{(\rho + r)\,\Gamma(\rho + r)}{\Gamma(\rho)\,\Gamma(r+1)}\, p^{\rho}(1-p)^{r} =$$

$$= (1-p)\rho + (1-p)\, M_{bn}(\tilde{r} \mid \rho, p),$$

de unde se obține

$$M_{bn}(\tilde{r} \mid \rho, p) = \frac{\rho(1-p)}{p}. \tag{14.4.6}$$

În cazul nostru, unde $p = \tau/(\tau + t)$, avem

$$M_{bn}\left(\tilde{r} \mid \rho,\ \frac{\tau}{\tau + t}\right) = \frac{\rho t}{\tau}. \tag{14.4.7}$$

Scanned with OKEN Scanner

Dispersia lui $\tilde{r} \in Bn(\rho, p)$ se obține de asemenea prin calcul direct. Avem mai întîi

$$M_{bn}(\tilde{r}^2 | \rho, p) - M_{bn}(\tilde{r} | \rho, p) =$$

$$= \sum_{r=0}^{\infty} (r^2 - r) \frac{\Gamma(\rho + r)}{\Gamma(\rho)\,\Gamma(r+1)} p^{\rho}(1-p)^r =$$

$$= \sum_{r=2}^{\infty} \frac{\Gamma(\rho + r)}{\Gamma(\rho)\,\Gamma(r-1)} p^{\rho}(1-p)^r = \sum_{r=0}^{\infty} \frac{\Gamma(\rho + r + 2)}{\Gamma(\rho)\,\Gamma(r+1)} p^{\rho}(1-p)^{r+2} =$$

$$= (1-p)^2 \sum_{r=0}^{\infty} \frac{(\rho + r + 1)\,(\rho + r)\,\Gamma(\rho + r)}{\Gamma(\rho)\,\Gamma(r+1)} p^{\rho}(1-p)^r =$$

$$= (1-p)^2 [\rho(\rho+1) + (2\rho+1)\, M_{bn}(\tilde{r} | \rho, p) + M_{bn}(\tilde{r}^2 | \rho, p)],$$

de unde se obține ușor

$$M_{bn}(\tilde{r}^2 | \rho, p) = \frac{\rho(1-\rho)\,[1 + \rho(1-p)]}{p^2}. \qquad (14.4.8)$$

Dispersia lui $\tilde{r} \in Bn(\rho, p)$ rezultă imediat din (14.4.7) și (14.4.8) și este dată de relația

$$V(\tilde{r} | \rho, p) = \frac{\rho(1-p)}{p^2}. \qquad (14.4.9)$$

În cazul nostru, unde $p = \tau/(\tau + t)$, obținem

$$V\left(\tilde{r} | \rho, \frac{\tau}{\tau + t}\right) = \frac{\rho t(\tau + t)}{\tau^2}. \qquad (14.4.10)$$

Dacă asociem acum procesului Poisson de generare de date un proces de oprire caracterizat prin faptul că observațiile durează pînă în momentul în care s-au produs exact r apariții ale fenomenului considerat (vezi § 12.7.4), atunci durata observațiilor $\tilde{t}$ are densitatea de probabilitate $f_r(t \mid r, \lambda)$. Dacă se utilizează pentru $\tilde{\lambda}$ o densitate de probabilitate apriori

$f_\gamma(\lambda|\rho,\tau)$ atunci densitatea de probabilitate a lui $\tilde{t}$ are valori nenule de forma

$$f(t)=\frac{f(t,\lambda)}{f(\lambda|t)}=\frac{f_\gamma(t|r,\lambda)\,f_\gamma(\lambda|\rho,\tau)}{f_\gamma(\lambda|\rho+r,\tau+t)}=$$

$$=\frac{\Gamma(\rho+r)}{\Gamma(\rho)\,\Gamma(r)}\cdot\frac{\lambda\tau}{\tau+t}\cdot\frac{e^{-\lambda t}e^{-\lambda\tau}(\lambda t)^{r-1}(\lambda\tau)^{\rho-1}}{e^{-\lambda(t+\tau)}[\lambda(t+\tau)]^{\rho+r-1}}=$$

$$=\frac{t^{r-1}\tau^{\rho}}{B(\rho,r)\,(t+\tau)^{\rho+r}}, \qquad (14.4.11)$$

ceea ce arată că $f(t)$ dat de (14.4.11) coincide cu densitatea de probabilitate $f_{\beta 2i}(t\mid r,\rho,\tau)$ a repartiției beta-2 inversă.

Utilizînd rezultatele (13.3.22) și (13.3.24), obținem

$$M_{\beta 2i}(\tilde{t}\,|\,r,\rho,\tau)=\frac{\tau r}{\rho-1}, \qquad (14.4.12)$$

cu condiția ca $\rho>1$ și

$$V_{\beta 2i}(\tilde{t}\,|\,r,\rho,\tau)=\frac{\tau^2 r(r+\rho-1)}{(\rho-1)^2(\rho-2)}, \qquad (14.4.13)$$

cu condiția ca $\rho>2$.

14.5. PROCESUL NORMAL UNIDIMENSIONAL

În § 11.6 am introdus procesul normal unidimensional $(\tilde{x}_i)$ de parametrii m și $\sigma^2>0$, pentru care am obținut funcțiile de verosimilitate corespunzătoare diverselor cazuri privind cunoașterea valorilor acestor parametrii (vezi § 11.6.1). Pentru același proces am obținut în § 13.9.1 repartițiile diverselor statistici care pot fi asociate acestui proces de generare de date, în cazul în care procesul de oprire asociat este neinformativ asupra parametrului în discuție.

În continuare vom prezenta familiile de repartiții conjugate cu procesul normal de generare de date de parametrii m și $\sigma^2 > 0$, analizînd separat cele trei cazuri care pot să apară (cazurile *A*, *B*, *C* menționate în § 11.6.1).

Teorema 14.5.1 (Cazul A). *Fie* $\tilde{x}_1, \ldots, \tilde{x}_n$ *o selecție de volum* n *din* $\tilde{x} \in N(m, \sigma^2)$, *unde* m *este necunoscut, iar* σ^2 *este cunoscut. Dacă se consideră pentru* $\tilde{m}$ *densitatea de probabilitate apriori* $f_N(m|m_0, \sigma^2/n_0)$, *atunci densitatea de probabilitate aposteriori a lui* $\tilde{m}$ *este* $f_N(m|m_1, \sigma^2/n_1)$, *unde*

$$m_1 = \frac{n\bar{x} + n_0 m_0}{n + n_0} \tag{14.5.1}$$

și

$$n_1 = n + n_0. \tag{14.5.2}$$

În plus, media de selecție $\tilde{\bar{x}}$ *are în aceste condiții densitatea de probabilitate* $f_N\left(\bar{x}|\, m_0, \frac{n + n_0}{n n_0}\sigma^2\right)$.

Demonstrație. În acest caz, conform teoremei 11.6.1, statistica $(\tilde{n}, \tilde{\bar{x}})$ este suficientă pentru m, iar funcția de verosimilitate $L(m|n, \bar{x}, \sigma^2)$ este de forma (11.6.12). Dacă densitatea de probabilitate apriori a lui $\tilde{m}$ este $f_N(m|m_0, \sigma^2/n_0)$, rezultă ușor că densitatea de probabilitate aposteriori a lui $\tilde{m}$ este proporțională cu produsul (vezi și formula (14.1.2))

$$f_N(m|\, m_0, \sigma^2/n_0)\, L(m|n, \bar{x}, \sigma^2),$$

adică este proporțională cu

$$\exp\left[-\frac{n_0(m - m_0)^2}{2\sigma^2} - \frac{n(\bar{x} - m)^2}{2\sigma^2}\right]. \tag{14.5.3}$$

Dar

$$\begin{aligned} n_0(m - m_0)^2 + n(\bar{x} - m)^2 = \\ = (n_0 + n)m^2 - 2(m_0 n_0 + n\bar{x})m + n_0 m_0^2 + n\bar{x}^2 = \\ = (n_0 + n)m^2 - 2(n_0 + n)m_1 m + (n_0 + n)m_1^2 - \\ - (n_0 + n)m_1^2 + n_0 m_0^2 + n\bar{x}^2 = n_1 m^2 - 2n_1 m_1 m + \\ + n_1 m_1^2 - n_1 m_1^2 + n_0 m_0^2 + n\bar{x}^2 = n_1(m - m_1)^2 + n_0 m_0^2 + \\ + n\bar{x}^2 - n_1 m_1^2. \end{aligned}$$

Scanned with OKEN Scanner

Rezultă atunci că expresia (14.5.3) devine

$$\exp\left[-\frac{n_1(m-m_1)^2}{2\sigma^2}\right]\exp\left[-\frac{1}{2\sigma^2}(n_0m_0^2+n\overline{x}^2-n_1m_1^2)\right]. \tag{14.5.4}$$

Cum al doilea factor din (14.5.4) nu depinde de m, rezultă că densitatea de probabilitate aposteriori a lui $\widetilde{m}$ trebuie să fie $f_N(m|m_1, \sigma^2/n_1)$, ceea ce demonstrează prima afirmație a teoremei.

Pentru a demonstra și a doua afirmație a teoremei, să observăm că din relația generală (14.1.3) rezultă că densitatea de probabilitate a lui $\widetilde{\overline{x}}$ este dată de relația

$$f(\overline{x}) = \frac{f(\overline{x}|m)\,f(m)}{f(m|\overline{x})} = \frac{f_N(\overline{x}|m, \sigma^2/n)\,f_N(m|m_0, \sigma^2/n_0)}{f_N(m|m_1, \sigma^2/n_1)}, \tag{14.5.5}$$

dacă ținem seama și de faptul că $\widetilde{\overline{x}} \in N(m, \sigma^2/n)$ pentru m și σ^2 dați.

Deoarece $f(\overline{x})$ conține partea independentă de m din produsul care apare în (14.5.5), rezultă că $f(\overline{x})$ este proporțional cu al doilea factor din expresia (14.5.4). Însă, punînd

$$y = n_0m_0^2 + n\overline{x}^2 - n_1m_1^2, \tag{14.5.6}$$

rezultă

$$(n_0+n)y = (n_0+n)n_0m_0^2 + (n_0+n)n\overline{x}^2 -$$

$$-(n_0+n)^2\frac{(n\overline{x}+n_0m_0)^2}{(n_0+n)^2} = (n_0+n)(n_0m_0^2+\overline{x}^2) -$$

$$-(n^2\overline{x}^2 + 2nn_0\overline{x}m_0 + n_0^2m_0^2) = nn_0(\overline{x}-m_0)^2,$$

ceea ce arată că al doilea factor din (14.5.4) se poate scrie de asemenea sub forma

$$\exp\left[-\frac{nn_0(\overline{x}-m)^2}{2\sigma^2(n+n_0)}\right],$$

și deci densitatea de probabilitate a lui $\widetilde{\overline{x}}$ este $f_N\left(\overline{x}|m_0, \frac{n+n_0}{nn_0}\sigma^2\right)$,

adică tocmai ceea ce trebuia demonstrat.

Scanned with OKEN Scanner

Observație. Dacă pentru $\tilde{m}$ adoptăm densitatea de probabilitate apriori $f_N(m|m_0, \sigma_0^2)$ în cazul cînd $\sigma^2 > 0$ este cunoscut, atunci rezultă că n_0 este *definit* de relația $\sigma^2 = n_0 \sigma_0^2$.

Este de asemenea instructiv să observăm că media aposteriori m_1 a lui $\tilde{m}$ se exprimă în funcție de media sa apriori m_0 și de media aritmetică a valorilor observate $\bar{x}$ prin relația (14.5.1), care poate fi scrisă și sub forma

$$m_1 = \frac{n}{n+n_0}\bar{x} + \frac{n_0}{n+n_0}m_0,$$

de unde se vede că m_1 este *media ponderată* a lui $\bar{x}$ și m_0, cu ponderile $n/(n+n_0)$, respectiv $n_0/(n+n_0)$. Cu cît n este mai mare, ponderea asociată a lui $\bar{x}$ se mărește.

În sfîrșit, să observăm că dispersia aposteriori a lui $\tilde{m}$ este $\sigma^2/(n+n_0)$ și descrește odată cu creșterea numărului observațiilor n.

Teorema 14.5.2 (Cazul B). *Fie* $\tilde{x}_1, \ldots, \tilde{x}_n$ *o selecție de volum* n *din* $\tilde{x} \in N(m, \sigma^2)$. *Dacă* $\tilde{\sigma}^2$ *are densitatea de probabilitate apriori* $f_{\gamma i}(\sigma^2 \mid \psi_0, \nu_0)$, *atunci densitatea de probabilitate aposteriori a lui* $\tilde{\sigma}^2$ *este* $f_{\gamma i}(\sigma^2|\psi_1, \nu_1)$, *unde*

$$\psi_1 = \frac{\psi_0\nu_0 + ns^2(m)}{n+\nu_0} \tag{14.5.7}$$

și

$$\nu_1 = n + \nu_0. \tag{14.5.8}$$

În plus, densitatea de probabilitate a lui $\tilde{s}^2(m)$ *pentru* m *și* n *dați este* $f_{\beta 2i}(s^2(m) \mid n/2, \nu_0/2, \psi_0\nu_0/n)$.

Demonstrație. Conform teoremei 11.6.2, funcția de verosimilitate pentru σ^2 este în acest caz de forma (11.6.17). Dacă densitatea de probabilitate apriori a lui $\tilde{\sigma}^2$ este $f_{\gamma i}(\sigma^2 \mid \psi_0, \nu_0)$, atunci produsul

$$L(\sigma^2 \mid n, x_1, \ldots, x_n, m)\, f_{\gamma i}(\sigma^2 \mid \psi_0, \nu_0)$$

este proporțional cu

$$(\sigma^2)^{-\frac{n+\nu_0+2}{2}} \exp\left[-\frac{1}{2\sigma^2}(ns^2(m) + \psi_0\nu_0)\right]$$

sau

$$(\sigma^2)^{-\frac{\nu_1+2}{2}} \exp\left(-\frac{\psi_1 \nu_1}{2\sigma^2}\right),$$

de unde rezultă imediat că densitatea de probabilitate aposteriori a lui $\tilde{\sigma}^2$ este $f_{\gamma i}(\sigma^2 \mid \psi_1, \nu_1)$.

Pentru a demonstra cea de a doua afirmație a teoremei, să observăm că pentru σ^2 dat, conform teoremei 13.9.1, densitatea de probabilitate a lui $\tilde{s}^2(m)$ este $f_{\gamma 2}(s^2(m) \mid 1/\sigma^2, n)$. Prin urmare densitatea de probabilitate a lui $\tilde{s}^2(m)$ se scrie sub forma

$$f(s^2(m)) = f(s^2(m) \mid \sigma^2) \frac{f(\sigma^2)}{f(\sigma^2 \mid s^2(m))} =$$

$$= f_{\gamma 2}(s^2(m) \mid 1/\sigma^2, n) \frac{f_{\gamma i}(\sigma^2 \mid \psi_0, \nu_0)}{f_{\gamma i}(\sigma^2 \mid \psi_1, \nu_1)}.$$

După calcule elementare obținem

$$f(s^2(m)) = \frac{(s^2(m))^{n/2-1} (\psi_0 \nu_0/n)^{\nu_0/2}}{B(n/2, \nu_0/2) (\psi_0 \nu_0/n + s^2(m))^{(n+\nu_0)/2}},$$

ceea ce arată că $\tilde{s}^2(m)$ aparține clasei $\beta 2i(n/2, \nu_0/2, \psi_0 \nu_0/n)$ și teorema este demonstrată.

Pentru cazul cînd nici unul dintre parametrii m și σ^2 nu este cunoscut vom considera drept densitate de probabilitate conjugată cu procesul normal de parametrii m și σ^2 așa-numita densitate de probabilitate *normală-gama inversă* notată prin $f_{N\gamma i}(m, \sigma^2 \mid m_0, n_0, \psi_0, \nu_0)$ și definită ca fiind egală cu produsul

$$f_N(m \mid m_0, \sigma^2/n_0) f_{\gamma i}(\sigma^2 \mid \psi_0, \nu_0), \qquad (14.5.9)$$

unde $m_0, n_0 > 0$, $\psi_0 > 0$ și $\nu_0 > 0$ sînt parametrii reali. Ca de obicei vom nota prin $N\gamma i(m_0, n_0, \psi_0, \nu_0)$ mulțimea vectorilor aleatori bidimensionali care au această repartiție.

Teorema 14.5.3. *Dacă* $(\tilde{m}, \tilde{\sigma}^2)$ *aparține clasei* $N\gamma i(m_0, n_0, \psi_0, \nu_0)$, *atunci* $\tilde{m}$ *aparține clasei* $S(m_0, \psi_0/n, \nu_0)$, *iar* $\tilde{\sigma}^2$ *aparține clasei* $\gamma i(\psi_0, \nu_0)$.

Demonstrație. Densitatea de probabilitate (marginală) a lui $\tilde{m}$ este deci de forma

$$f(m) = \int_0^\infty f_N(m \mid m_0,\ \sigma^2/n_0) f_{\gamma i}(\sigma^2 \mid \psi_0,\ \nu_0)\, d\sigma^2.$$

Deoarece $1/\tilde{\sigma}^2$ aparține clasei $\gamma 2(\psi_0,\ \nu_0)$, rezultă că

$$f(m) = \int_0^\infty f_N\left(m \mid m_0,\ \frac{1}{x n_0}\right) f_{\gamma 2}(x \mid \psi_0,\ \nu_0)\, dx,$$

de unde, după calcule simple, se obține

$$f(m) = f_S(m \mid m_0,\ \psi_0/n_0,\ \nu_0).$$

Faptul că $\tilde{\sigma}^2$ aparține clasei $\gamma i(\psi_0,\ \nu_0)$ este evident.

Teorema 14.5.4. *Fie* $\tilde{x}_1, \ldots, \tilde{x}_n$ *o selecție de volum* n *din* $\tilde{x} \in N(m, \sigma^2)$, *unde* m *și* σ^2 *sînt necunoscuți. Dacă* $(\tilde{m}, \tilde{\sigma}^2)$ *are densitatea de probabilitate apriori* $f_{N\gamma i}(m, \sigma^2 \mid m_0, n_0, \psi_0, \nu_0)$, *atunci densitatea de probabilitate aposteriori a lui* $(\tilde{m}, \tilde{\sigma}^2)$ *este* $f_{N\gamma i}(m, \sigma^2 \mid m_1, n_1, \psi_1, \nu_1)$, *unde*

$$m_1 = \frac{n\bar{x} + n_0 m_0}{n + n_0}, \tag{14.5.10}$$

$$n_1 = n + n_0, \tag{14.5.11}$$

$$\nu_1 = n + \nu_0 \tag{14.5.12}$$

și

$$\psi_1 = \frac{1}{\nu_1}\left[\psi_0 \nu_0 + (n-1)s^2 + \frac{n n_0}{n_1}(\bar{x} - m_0)^2\right]. \tag{14.5.13}$$

Demonstrație. Să observăm mai întîi că funcția de verosimilitate $L(m, \sigma^2 \mid n, x_1, \ldots, x_n)$ este, conform teoremei 11.6.3, proporțională cu produsul

$$(\sigma^2)^{-n/2} \exp\left[-\frac{1}{2\sigma^2}(n-1)s^2\right] \exp\left[-\frac{1}{2\sigma^2} n(\bar{x} - m)^2\right].$$

Rezultă atunci că produsul

$$L(m, \sigma^2 \mid n, x_1, \ldots, x_n)\, f_{N\gamma i}(m, \sigma^2 \mid m_0, n_0, \psi_0, \nu_0) \quad (14.5.14)$$

este proporțional cu produsul următoarelor două expresii:

$$(\sigma^2)^{-1/2} \exp\left[-\frac{n_0}{2\sigma^2}(m - m_0)^2\right] \exp\left[-\frac{n}{2\sigma^2}(\overline{x} - m)^2\right] \quad (14.5.15)$$

și

$$(\sigma^2)^{-\frac{n+\nu_0}{2}-1} \exp\left[-\frac{1}{2\sigma^2}(n-1)s^2 - \frac{\psi_0 \nu_0}{2\sigma^2}\right]. \quad (14.5.16)$$

Să observăm acum că, abstracție făcînd de factorul $(\sigma^2)^{-1/2}$, expresia (14.5.15) coincide cu (14.5.3). Ținînd seama de relațiile obținute în demonstrația teoremei 14.5.1, rezultă că putem scrie expresia (14.5.15) sub forma

$$(\sigma^2)^{-1/2} \exp\left[-\frac{n_1(m - m_1)^2}{2\sigma^2}\right] \exp\left[-\frac{nn_0(\overline{x} - m)}{2\sigma^2(n + n_0)}\right]. \quad (14.5.17)$$

Utilizînd relațiile (14.5.16) și (14.5.17) rezultă că produsul (14.5.14) este proporțional cu

$$\left\{(\sigma^2)^{-1/2} \exp\left[-\frac{n_1(m - m_1)^2}{2\sigma^2}\right]\right\} \cdot \left\{(\sigma^2)^{\nu_1/2-1} \exp\left[-\frac{\psi_1 \nu_1}{2\sigma^2}\right]\right\}. \quad (14.5.18)$$

Cum prima acoladă din (14.5.18) este proporțională cu $f_N(m \mid m_1, \sigma^2/n_1)$, iar a doua acoladă din (14.5.18) este proporțională cu $f_{\gamma i}(\sigma^2 \mid \psi_1, \nu_1)$, rezultă că produsul (14.5.14) este proporțional cu

$$f_N(m \mid m_1, \sigma^2/n_1)\, f_{\gamma i}(\sigma^2 \mid \psi_1, \nu_1),$$

de unde se obține imediat afirmația din teoremă.

Teorema 14.5.5 (Cazul C). *Fie* $\tilde{x}_1, \ldots, \tilde{x}_n$ *o selecție de volum* n *din* $\tilde{x} \in N(m, \sigma^2)$, *unde* m *și* σ^2 *sînt necunoscuți.*

Dacă $(\tilde{m}, \tilde{\sigma}^2)$ *are densitatea de probabilitate apriori* $f_{N\gamma i}(m, \sigma^2 \mid m_0, n_0, \psi_0, \nu_0)$, *atunci:*

(a) Densitatea de probabilitate condiționată $f(\bar{x} \mid s^2)$ *este dată de relația*

$$f(\bar{x} \mid s^2) = f_S\left(\bar{x} \mid m_0, \frac{n_1}{nn_0}\psi^*, \nu_1 - 1\right), \qquad (14.5.19)$$

unde

$$\psi^* = \frac{\nu_0}{\nu_1 - 1}\psi_0 + \frac{n-1}{\nu_1 - 1}s^2; \qquad (14.5.20)$$

(b) Densitatea de probabilitate $f(\bar{x})$ *este dată de relația*

$$f(\bar{x}) = f_S\left(\bar{x} \mid m_0, \frac{n_1}{nn_0}\psi_0, \nu_0\right); \qquad (14.5.21)$$

(c) Densitatea de probabilitate condiționată $f(s^2 \mid \bar{x})$ *este dată de relația*

$$f(s^2 \mid \bar{x}) = f_{\beta 2i}(s^2 \mid (n-1)/2, (\nu_0 + 1)/2, b), \qquad (14.5.22)$$

unde

$$b = \frac{1}{n-1}\left[\psi_0 \nu_0 + \frac{n_1}{nn_0}(\bar{x} - m_0)^2\right]; \qquad (14.5.23)$$

(d) Densitatea de probabilitate $f(s^2)$ *este dată de relația*

$$f(s^2) = f_{\beta 2i}(s^2 \mid (n-1)/2, (\nu_0 + 1)/2, \psi_0 \nu_0/(n-1)). \qquad (14.5.24)$$

Demonstrație. Pentru demonstrarea acestor afirmații se utilizează aceeași metodă care a fost folosită și în celelalte demonstrații date mai sus. Din cauza lungimii calculelor elementare necesare, precum și fiindcă vom da demonstrațiile corespunzătoare pentru cazul multidimensional, omitem această demonstrație.

Scanned with OKEN Scanner

14.6. PROCESUL NORMAL MULTIDIMENSIONAL

Vom prezenta în acest paragraf rezultate analoage celor date în § 14.5 pentru procesul normal unidimensional.

În § 11.7 am introdus procesul normal k—dimensional $(\tilde{\mathbf{x}}_t)$ de parametrii $\mathbf{m}$ și $\boldsymbol{\delta}^2$, pentru care am obținut funcțiile de verosimilitate corespunzătoare diverselor cazuri privind cunoașterea acestor parametrii (vezi § 11.7.1). Pentru același proces am obținut în § 13.9.2 repartițiile diverselor statistici care pot fi asociate acestui proces de generare de date, în cazul în care procesul de oprire este neinformativ asupra parametrului în discuție.

În analogie cu rezultatele date în § 14.5 pentru procesul normal unidimensional, în cele ce urmează vom prezenta familiile de repartiții conjugate cu procesul normal k—dimensional de generare de date de parametrii $\mathbf{m}$ și $\boldsymbol{\delta}^2$, analizînd separat diverse cazuri interesante privind cunoașterea acestor parametrii (vezi cazurile A, B, ..., E menționate în § 11.7.1). Ordinea de prezentare a rezultatelor pentru aceste cazuri ține seama de complexitatea analizei lor.

Teorema 14.6.1 (Cazul A). *Fie* $\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_n$ *o selecție de volum* n *din* $\tilde{\mathbf{x}} \in N_k(\mathbf{m}, \boldsymbol{\delta}^2)$ *unde* $\mathbf{m}$ *este necunoscut, iar* $\boldsymbol{\delta}^2$ *este cunoscut.*

Dacă $\tilde{\mathbf{m}}$ *are densitatea de probabilitate apriori* $f_N(\mathbf{m}|\mathbf{m}_0, \boldsymbol{\delta}_0^2)$ *atunci densitatea sa de probabilitate aposteriori este* $f_N(\mathbf{m}|\mathbf{m}_1, \boldsymbol{\delta}_1^2)$, *unde*

$$\boldsymbol{\delta}_1^2 = (\boldsymbol{\delta}_0^{-2} + n\,\boldsymbol{\delta}^{-2})^{-1} \tag{14.6.1}$$

și

$$\mathbf{m}_1 = \boldsymbol{\delta}_1^2(\boldsymbol{\delta}_0^{-2}\mathbf{m}_0 + n\,\boldsymbol{\delta}^{2}\bar{\mathbf{x}}). \tag{14.6.2}$$

În plus, media de selecție $\tilde{\bar{\mathbf{x}}}$ *are densitatea de probabilitate*

$$f_N\left(\bar{\mathbf{x}} \mid \mathbf{m}_0,\ \boldsymbol{\delta}_0^2 + \frac{1}{n}\,\boldsymbol{\delta}^2\right).$$

Demonstrație. Conform teoremei 11.7.1, funcția de verosimilitate $L(\mathbf{m} \mid n, \bar{\mathbf{x}}, \boldsymbol{\delta}^2)$ este în acest caz proporțională cu

$$\exp\left[-\frac{n}{2}\,(\bar{\mathbf{x}} - \mathbf{m})'\,\boldsymbol{\delta}^{-2}(\bar{\mathbf{x}} - \mathbf{m})\right].$$

Dacă $\tilde{\mathrm{m}}$ are densitatea de probabilitate apriori $f_N(\mathrm{m} \mid \mathrm{m}_0, \delta_0^2)$, rezultă că densitatea de probabilitate aposteriori a lui $\tilde{\mathrm{m}}$ este proporțională cu produsul

$$L(\mathrm{m} \mid n, \bar{\mathrm{x}}, \delta^2)\, f_N(\mathrm{m} \mid \mathrm{m}_0, \delta_0^2), \tag{14.6.3}$$

adică este proporțională cu

$$\exp\left[-\frac{n}{2}(\bar{\mathrm{x}}-\mathrm{m})'\,\delta^{-2}(\bar{\mathrm{x}}-\mathrm{m})\right]\exp\left[-\frac{1}{2}(\mathrm{m}-\mathrm{m}_0)'\,\delta_0^{-2}(\mathrm{m}-\mathrm{m}_0)\right],$$

expresie echivalentă cu

$$\exp\left\{-\frac{1}{2}\left[n(\bar{\mathrm{x}}-\mathrm{m})'\,\delta^{-2}(\bar{\mathrm{x}}-\mathrm{m})+(\mathrm{m}-\mathrm{m}_0)'\,\delta_0^{-2}(\mathrm{m}-\mathrm{m}_0)\right]\right\}.$$

Însă

$$\begin{aligned}
&n(\bar{\mathrm{x}}-\mathrm{m})'\,\delta^{-2}(\bar{\mathrm{x}}-\mathrm{m})+(\mathrm{m}-\mathrm{m}_0)'\,\delta_0^{-2}(\mathrm{m}-\mathrm{m}_0)=\\
&=\bar{\mathrm{x}}'(n\,\delta^{-2})\,\bar{\mathrm{x}}+\mathrm{m}_0'\,\delta_0^{-2}\mathrm{m}_0-\mathrm{m}'(\delta_0^{-2}\mathrm{m}_0+n\,\delta^{-2}\bar{\mathrm{x}})-\\
&\quad-(\mathrm{m}_0'\,\delta_0^{-2}+\bar{\mathrm{x}}'\,n\,\delta^{-2})\,\mathrm{m}+\mathrm{m}'(n\,\delta^{-2}+\delta_0^{-2})=\\
&=\bar{\mathrm{x}}'(\mathrm{n}\,\delta^{-2})\bar{\mathrm{x}}+\mathrm{m}_0'\,\delta_0^{-2}\mathrm{m}_0-\mathrm{m}'\,\delta_1^{-2}\mathrm{m}_1-(\delta_1^{-2}\mathrm{m}_1)'\mathrm{m}+\mathrm{m}'\,\delta_1^{-2}\mathrm{m}=\\
&=(\mathrm{m}-\mathrm{m}_1)'\,\delta_1^{-2}(\mathrm{m}-\mathrm{m}_1)-\mathrm{m}_1'\,\delta_1^{-2}\mathrm{m}_1+\mathrm{m}_0'\,\delta_0^{-2}\mathrm{m}_0+\bar{\mathrm{x}}'(n\,\delta^{-2})\,\bar{\mathrm{x}},
\end{aligned}$$

ceea ce arată că produsul (14.6.3) este proporțional cu

$$k_0\exp\left[-\frac{1}{2}(\mathrm{m}-\mathrm{m}_1)'\,\delta_1^{-2}(\mathrm{m}-\mathrm{m}_1)\right], \tag{14.6.4}$$

unde

$$k_0=\exp\left\{-\frac{1}{2}\left[\bar{\mathrm{x}}'(n\delta^{-2})\bar{\mathrm{x}}+\mathrm{m}_0'\delta_0^{-2}\mathrm{m}_0-\mathrm{m}_1'\delta_1^{-2}\mathrm{m}_1\right]\right\} \tag{14.6.5}$$

nu depinde de m.

Din (14.6.3) și (14.6.4) rezultă că densitatea de probabilitate aposteriori a lui $\tilde{\mathrm{m}}$ trebuie să fie $f_N(\mathrm{m} \mid \mathrm{m}_1, \delta_1^2)$, ceea ce demonstrează prima afirmație a teoremei.

Scanned with OKEN Scanner

Pentru a demonstra cea de a doua afirmație, să observăm că funcția caracteristică a lui $\tilde{\tilde{\mathbf{x}}}$ poate fi scrisă sub forma următoare:

$$\varphi(\mathbf{t}) = M(e^{i\mathbf{t}'\tilde{\tilde{\mathbf{x}}}}) = M(M(e^{i\mathbf{t}'\tilde{\tilde{\mathbf{x}}}} \mid \tilde{\mathbf{m}})).$$

Deoarece, pentru $\mathbf{m}$ dat, $\tilde{\tilde{\mathbf{x}}}$ aparține clasei $N_k\left(\mathbf{m}, \frac{1}{n}\,\boldsymbol{\delta}^2\right)$, conform formulei (10.3.35) rezultă că

$$M(e^{i\mathbf{t}'\tilde{\tilde{\mathbf{x}}}} \mid \tilde{\mathbf{m}}) = \exp\left(i\mathbf{t}'\tilde{\mathbf{m}} - \frac{1}{2n}\,\mathbf{t}'\boldsymbol{\delta}^2\mathbf{t}\right).$$

Cum $\tilde{\mathbf{m}}$ are densitatea de probabilitate $f_N(\mathbf{m} \mid \mathbf{m}_0, \boldsymbol{\delta}_0^2)$, rezultă că

$$M(e^{i\mathbf{t}'\tilde{\mathbf{m}}}) = \exp\left(i\mathbf{t}'\mathbf{m}_0 - \frac{1}{2}\,\mathbf{t}'\boldsymbol{\delta}_0^2\mathbf{t}\right)$$

și deci

$$\varphi(\mathbf{t}) = \exp\left[i\mathbf{t}'\mathbf{m}_0 - \frac{1}{2}\,\mathbf{t}'\left(\frac{1}{n}\,\boldsymbol{\delta}^2 + \boldsymbol{\delta}_0^2\right)\mathbf{t}\right],$$

de unde rezultă că $\tilde{\tilde{\mathbf{x}}}$ aparține clasei $N\left(\mathbf{m}_0, \boldsymbol{\delta}_0^2 + \frac{1}{n}\,\boldsymbol{\delta}^2\right)$, adică tocmai ceea ce trebuia demonstrat.

Teorema 14.6.2 (Cazul D). *Fie* $\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_n$ *o selecție de volum* $\mathbf{n}$ *din* $\tilde{\mathbf{x}} \in N_k(\mathbf{m}, \boldsymbol{\delta}^2)$, *unde* $\mathbf{m}$ *este cunoscut, iar* $\boldsymbol{\delta}^2$ *este necunoscut.*

Dacă se consideră pentru $\tilde{\boldsymbol{\delta}}^2$ *densitatea de probabilitate apriori* $f_{Wi}(\boldsymbol{\delta}^2 \mid \psi_0, \nu_0)$, *atunci densitatea de probabilitate aposteriori a lui* $\tilde{\boldsymbol{\delta}}^2$ *este* $f_{Wi}(\boldsymbol{\delta}^2 \mid \psi_1, \nu_1)$, *unde*

$$\psi_1 = \psi_0 + n\,\mathbf{s}^2(\mathbf{m}) \tag{14.6.6}$$

și

$$\nu_1 = n + \nu_0. \tag{14.6.7}$$

În plus, densitatea de probabilitate pentru $\tilde{\mathbf{s}}^2(\mathbf{m})$ *este* $f_{\beta 2i}(\mathbf{s}^2(\mathbf{m}) \mid \psi_0, \nu_0, n, n)$.

Demonstrație. În acest caz, conform teoremei 11.7.4, statistica $(\tilde{n}, \tilde{s}^2(m))$ este suficientă pentru δ^2, iar funcția de verosimilitate $L(\delta^2 \mid n, s^2(m))$ este proporțională cu

$$(\det \delta^2)^{-n/2} \exp\left[-\frac{n}{2} \operatorname{ur}(\delta^{-2} s^2(m)) \right].$$

Dacă $\tilde{\delta}^2$ are densitatea de probabilitate apriori $f_{Wi}(\delta^2 \mid \psi_0, \nu_0)$, atunci densitatea de probabilitate aposteriori a lui $\tilde{\delta}^2$ este proporțională cu produsul

$$L(\delta^2 \mid n, s^2(m))\, f_{Wi}(\delta^2 \mid \psi_0, \nu_0),$$

adică este proporțională cu

$$(\det \delta^2)^{-\frac{2n+\nu_0+1}{2}} \exp\left[-\frac{n}{2} \operatorname{ur}(\delta^{-2} s^2(m)) \right] \exp\left[-\frac{1}{2} \operatorname{ur}(\delta^{-2}\psi_0) \right]$$

sau

$$(\det \delta^2)^{-\frac{n+\nu_1+1}{2}} \exp\left[-\frac{1}{2} \operatorname{ur}(\delta^{-2}\psi_1) \right],$$

de unde decurge imediat prima afirmație a teoremei.

Pentru a demonstra cea de a doua afirmație, să observăm mai întîi că pentru δ^2 dat, conform teoremei 13.9.3, statistica $\tilde{s}^2(m)$ aparține clasei $W(n\,\delta^{-2}, n)$. Cum $\tilde{\delta}^2$ aparține clasei $Wi(\psi_0, \nu_0)$, rezultă că $(1/n)\tilde{\delta}^2$ aparține clasei $Wi((1/n)\psi_0, \nu_0)$ și deci, conform teoremei 13.8.1, rezultă că $\tilde{s}^2(m)$ are densitatea de probabilitate $f_{\beta 2i}(s^2(m) \mid (1/n)\,\psi_0, \nu_0, n, 1)$ sau, echivalent, $f_{\beta 2i}(s^2(m) \mid \psi_0, \nu_0, n, n)$, ceea ce încheie demonstrația teoremei.

Teorema 14.6.3 (Cazul C). *Fie* $\tilde{x}_1, \ldots, \tilde{x}_n$ *o selecție de volum* n *din* $\tilde{x} \in N_k(m, \delta^2)$, *unde* m *și* δ^2_* *sînt cunoscuți, iar* σ^2_* *este necunoscut* *.

---

* Vezi § 11.7.1 pentru definițiile lui δ^2_* și σ^2_*.

Dacă $\tilde{\sigma}_*^2$ *are densitatea de probabilitate apriori* $f_{\gamma i}(\sigma_*^2 \mid \psi_0, \nu_0)$, *atunci densitatea de probabilitate aposteriori a lui* $\tilde{\sigma}_*^2$ *este* $f_{\gamma i}(\sigma_*^2 \mid \psi_1, \nu_1)$, *unde*

$$\psi_1 = \frac{\nu_0}{\nu_1}\psi_0 + \frac{kn}{\nu_1}s_*^2(m) \tag{14.6.8}$$

și

$$\nu_1 = kn + \nu_0. \tag{14.6.9}$$

În plus, densitatea de probabilitate a lui $\tilde{s}_*^2(m)$ *este* $f_{\beta 2i}(s_*^2(m) \mid kn/2, \nu_0/2, \psi_0\nu_0/kn)$.

Demonstrație. Din teorema 11.7.3 rezultă în acest caz că statistica $(\tilde{n}, \tilde{s}_*^2(m))$ este suficientă pentru σ_*^2, iar funcția de verosimilitate $L(\sigma_*^2 | n, s_*^2(m))$ este proporțională cu

$$(\sigma_*^2)^{-kn/2}\exp\left[-\frac{kn}{2}s_*^2(m)/\sigma_*^2\right].$$

Dacă se adoptă pentru $\tilde{\sigma}_*^2$ densitatea de probabilitate apriori $f_{\gamma i}(\sigma_*^2 \mid \psi_0, \nu_0)$, rezultă că densitatea de probabilitate aposteriori a lui $\tilde{\sigma}_*^2$ este proporțională cu produsul

$$L(\sigma_*^2 \mid n, s_*^2(m))\ f_{\gamma i}(\sigma_*^2 \mid \psi_0, \nu_0),$$

adică este proporțională cu produsul

$$(\sigma_*^2)^{-\frac{kn+\nu_0+2}{2}}\exp\left[-\frac{kn}{2}s_*^2(m)/\sigma_*^2\right]\exp\left[-\frac{\psi_0\nu_0}{2\sigma_*^2}\right]$$

sau

$$(\sigma_*^2)^{-\frac{\nu_1}{2}-1}\exp\left[-\frac{1}{2}(kns_*^2(m)+\psi_0\nu_0)/\sigma_*^2\right]$$

sau încă

$$(\sigma_*^2)^{-\frac{\nu_1}{2}-1}\exp\left[-\frac{\psi_1\nu_1}{2\sigma_*^2}\right].$$

Scanned with OKEN Scanner

Din ultima formă se deduce imediat că densitatea de probabilitate aposteriori a lui $\tilde{\sigma}_*^2$ trebuie să fie $f_{\gamma i}(\sigma_*^2 \mid \psi_1, \nu_1)$.

Pentru a demonstra cea de a doua afirmație a teoremei, să observăm mai întîi că pentru σ_*^2 dat, conform teoremei 13.9.3, densitatea de probabilitate a lui $\tilde{s}_*^2(\mathrm{m})$ este $f_{\gamma 2}(s_*^2(\mathrm{m}) \mid 1/\sigma_*^2, kn)$. Rezultă atunci că

$$f(s_*^2(\mathrm{m})) = \frac{f(s_*^2(\mathrm{m}) \mid \sigma_*^2)\, f(\sigma_*^2)}{f(\sigma_*^2 \mid s_*^2(\mathrm{m}))} =$$

$$= \frac{f_{\gamma 2}(s_*^2(\mathrm{m}) \mid 1/\sigma_*^2, kn)\, f_{\gamma i}(\sigma_*^2 \mid \psi_0, \nu_0)}{f_{\gamma i}(\sigma_*^2 \mid \psi_1, \nu_1)},$$

de unde se obține ușor

$$f(s_*^2(\mathrm{m})) = f_{\beta 2i}(s_*^2(\mathrm{m}) \mid kn/2, \nu_0/2, \psi_0, \nu_0/kn),$$

ceea ce încheie demonstrația teoremei.

Ultimele două cazuri care mai rămîn de analizat (cazurile *B* și *E* menționate în § 11.7.1) presupun calcule mai complexe.

Vom începe analiza cu cazul *E*, cînd $\mathbf{m}$ și $\boldsymbol{\delta}^2$ sînt necunoscuți. În acest caz, densitatea de probabilitate conjugată pentru $(\tilde{\mathbf{m}}, \tilde{\boldsymbol{\delta}}^2)$ se dovedește a fi așa-numita densitate de probabilitate *normală-Wishart inversă*, notată prin $f_{NWi}(\mathbf{m}, \boldsymbol{\delta}^2 \mid \mathbf{m}_0, n_0, \boldsymbol{\Psi}_0, \nu_0)$ și definită ca fiind egală cu produsul

$$f_N\left(\mathbf{m}, \mid \mathbf{m}_0, \frac{1}{n_0}\boldsymbol{\delta}^2\right) f_{Wi}(\boldsymbol{\delta}^2 \mid \boldsymbol{\Psi}_0, \nu_0) \qquad (14.6.10)$$

pentru $\boldsymbol{\delta}^2$ simetrică și pozitiv definită și avînd valoarea zero în caz contrar.

Teorema 14.6.4 (Cazul E). *Dacă* $(\tilde{\mathbf{m}}, \tilde{\boldsymbol{\delta}}^2)$ *are densitatea de probabilitate* $\mathrm{f_{NWi}}(\mathbf{m}, \boldsymbol{\delta}^2 \mid \mathbf{m}_0, \mathrm{n}_0, \boldsymbol{\Psi}_0, \nu_0)$, *atunci densitatea de probabilitate (marginală) a lui* $\tilde{\mathbf{m}}$ *este*

$$\mathrm{f_S}\left(\mathbf{m} \mid \mathbf{m}_0, \frac{1}{\mathrm{n}_0(\nu_0 - \mathrm{k} + 1)}\boldsymbol{\Psi}_0, \nu_0 - k + 1\right),$$

iar densitatea de probabilitate (marginală) a lui $\tilde{\boldsymbol{\delta}}^2$ *este* $\mathrm{f_{Wi}}(\boldsymbol{\delta}^2 \mid \boldsymbol{\Psi}_0, \nu_0)$.

Scanned with OKEN Scanner

Demonstrație. Să observăm mai întîi că putem scrie

$$f_{NWi}(\mathbf{m}, \delta^2 \mid \mathbf{m}_0, n_0, \psi_0, \nu_0) =$$

$$= (2\pi)^{-\frac{k}{2}} \left(\det \frac{1}{n_0}\delta^2\right)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\mathbf{m}-\mathbf{m}_0)'(n_0\delta^{-2})(\mathbf{m}-\mathbf{m}_0)\right] \times$$

$$\times \frac{(\det \psi_0)^{\frac{\nu_0}{2}} (\det \delta^{-2})^{\frac{\nu_0+k+1}{2}} \exp\left[-\frac{1}{2}\operatorname{ur}(\delta^{-2}\psi_0)\right]}{\pi^{\frac{k(k-1)}{4}} 2^{\frac{k\nu_0}{2}} \Gamma\left(\frac{\nu_0}{2}\right)\ldots\Gamma\left(\frac{\nu_0-k+1}{2}\right)} =$$

$$= \frac{n_0^{\frac{k}{2}} (\det \psi_0)^{\frac{\nu_0}{2}} (\det \delta^{-2})^{\frac{\nu+k+1}{2}} \exp\left[-\frac{1}{2}\operatorname{ur}(\delta^{-2}\psi)\right]}{\pi^{\frac{k(k+1)}{4}} 2^{\frac{k\nu}{2}} \Gamma\left(\frac{\nu_0}{2}\right)\ldots\Gamma\left(\frac{\nu_0-k+1}{2}\right)},$$

unde

$$\psi = n_0(\mathbf{m}-\mathbf{m}_0)(\mathbf{m}-\mathbf{m}_0)' + \psi_0$$

și

$$\nu = \nu_0 + 1.$$

Se vede ușor că

$$f_{NWi}(\mathbf{m}, \delta^2 \mid \mathbf{m}_0, n_0, \psi_0, \nu_0) =$$

$$= \frac{n_0^{\frac{k}{2}} (\det \psi_0)^{\frac{\nu_0}{2}} \Gamma\left(\frac{\nu_0+1}{2}\right)}{\pi^{\frac{k}{2}} \Gamma\left(\frac{\nu_0-k+1}{2}\right) (\det \psi)^{\frac{\nu_0+1}{2}}} f_{Wi}(\delta^2 \mid \psi;\ \nu).$$

Integrînd în raport cu δ^2, rezultă că densitatea de probabilitate a lui $\tilde{m}$ este de forma

$$k_0 \left\{\det\left[(\mathbf{m}-\mathbf{m}_0)(\mathbf{m}-\mathbf{m}_0)' + \frac{1}{n_0}\psi_0\right]\right\}^{-\frac{\nu_0+1}{2}},$$

unde k_0 este un factor care nu depinde de m.

O teoremă de algebră liniară afirmă că, dacă A este o matrice de ordinul k nesingulară, iar $\mathbf{x} \in R^k$, atunci

$$\det(\mathbf{A} + \mathbf{x}\mathbf{x}') = (\det \mathbf{A})(1 + \mathbf{x}'\mathbf{A}^{-1}\mathbf{x}). \qquad (14.6.11)$$

Aplicînd această teoremă pentru $\mathbf{A} = \frac{1}{n_0}\psi_0$ și $\mathbf{x} = \mathbf{m} - \mathbf{m}_0$, rezultă că

$$\det\left[(\mathbf{m} - \mathbf{m}_0)(\mathbf{m} - \mathbf{m}_0)' + \frac{1}{n_0}\psi_0\right] =$$

$$= \left[\det\left(\frac{1}{n_0}\psi_0\right)\right][1 + (\mathbf{m} - \mathbf{m}_0)'(n_0\psi_0)^{-1}(\mathbf{m} - \mathbf{m}_0)].$$

Prin urmare densitatea de probabilitate a lui $\tilde{\mathbf{m}}$ este de forma

$$k_0'[1 + n_0(\mathbf{m} - \mathbf{m}_0)'\psi_0^{-1}(\mathbf{m} - \mathbf{m}_0)]^{-\frac{\nu_0 + 1}{2}},$$

unde k_0' este un factor care nu depinde de $\mathbf{m}$. Prima afirmație a teoremei este astfel demonstrată.

A doua afirmație a teoremei este evidentă.

Teorema 14.6.5 (Cazul E). *Fie* $\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_n$ *o selecție de volum* $\mathbf{n}$ *din* $\tilde{\mathbf{x}} \in N_k(\mathbf{m}, \boldsymbol{\delta}^2)$, *unde* $\mathbf{m}$ *și* $\boldsymbol{\delta}^2$ *sînt necunoscuți.*

Dacă se consideră pentru $(\tilde{\mathbf{m}}, \tilde{\boldsymbol{\delta}}^2)$ *densitatea de probabilitate apriori* $f_{NWi}(\mathbf{m}, \boldsymbol{\delta}^2 \mid \mathbf{m}_0, n_0, \psi_0, \nu_0)$, *atunci densitatea de probabilitate aposteriori pentru* $(\tilde{\mathbf{m}}, \tilde{\boldsymbol{\delta}}^2)$ *este* $f_{NWi}(\mathbf{m}, \boldsymbol{\delta}^2 \mid \mathbf{m}_1, \mathbf{n}_1, \psi_1, \nu_1)$, *unde*

$$\mathbf{n}_1 = \mathbf{n} + \mathbf{n}_0, \qquad (14.6.12)$$

$$\nu_1 = \mathbf{n} + \nu_0, \qquad (14.6.13)$$

$$\mathbf{m}_1 = \frac{1}{\mathbf{n}_1}(\mathbf{n}_0\mathbf{m}_0 + \mathbf{n}\bar{\mathbf{x}}) \qquad (14.6.14)$$

și

$$\psi_1 = \psi_0 + (\mathbf{n} - 1)\mathbf{s}^2 + \frac{\mathbf{n}_0\mathbf{n}}{\mathbf{n}_1}(\bar{\mathbf{x}} - \mathbf{m}_0)(\bar{\mathbf{x}} - \mathbf{m}_0)'. \qquad (14.6.15)$$

Scanned with OKEN Scanner

Demonstrație. În acest caz, conform teoremei 11.7.5, statistica $(\tilde{n}, \tilde{\bar{\mathbf{x}}}, \tilde{\mathbf{s}}^2)$ este suficientă pentru $(\mathbf{m}, \boldsymbol{\delta}^2)$, iar funcția de verosimilitate are forma

$$L(\mathbf{m}, \boldsymbol{\delta}^2 \mid n, \bar{\mathbf{x}}, \mathbf{s}^2) =$$

$$= k_0(\det \boldsymbol{\delta}^2)^{-n/2} \exp\left[-\frac{n-1}{2} \operatorname{ur}(\boldsymbol{\delta}^{-2}\mathbf{s}^2)\right] \times$$

$$\times \exp\left[-\frac{n}{2}(\bar{\mathbf{x}} - \mathbf{m})' \boldsymbol{\delta}^{-2}(\bar{\mathbf{x}} - \mathbf{m})\right],$$

unde

$$k_0 = (2\pi)^{-kn/2}.$$

Dacă se consideră pentru $(\tilde{\mathbf{m}}, \tilde{\boldsymbol{\delta}}^2)$ densitatea de probabilitate apriori $f_{NWi}(\mathbf{m}, \boldsymbol{\delta}^2 \mid \mathbf{m}_0, n_0, \boldsymbol{\psi}_0, \nu_0)$, atunci rezultă că densitatea de probabilitate aposteriori a lui $(\tilde{\mathbf{m}}, \tilde{\boldsymbol{\delta}}^2)$ este proporțională cu produsul

$$L(\mathbf{m}, \boldsymbol{\delta}^2 \mid n, \bar{\mathbf{x}}, \mathbf{s}^2)\, f_{NWi}(\mathbf{m}, \boldsymbol{\delta}^2 \mid \mathbf{m}_0, n_0, \boldsymbol{\psi}_0, \nu_0),$$

adică este proporțională cu produsul

$$(\det \boldsymbol{\delta}^2)^{-k/2} (\det \boldsymbol{\delta}^2)^{-\frac{\nu_1 + k + 1}{2}} \times$$

$$\times \exp\left\{-\frac{1}{2}\left[\operatorname{ur} \boldsymbol{\delta}^{-2}(\boldsymbol{\psi}_0 + (n-1)\mathbf{s}^2 + \boldsymbol{\psi})\right]\right\}, \qquad (14.6.16)$$

unde

$$\boldsymbol{\psi} = n(\bar{\mathbf{x}} - \mathbf{m})(\bar{\mathbf{x}} - \mathbf{m})' + n_0(\mathbf{m} - \mathbf{m}_0)(\mathbf{m} - \mathbf{m}_0)' \qquad (14.6.17)$$

Un calcul simplu ne conduce la concluzia că statistica $\boldsymbol{\psi}$ definită de (14.6.17) poate fi scrisă sub forma următoare

$$\boldsymbol{\psi} = n_1(\mathbf{m} - \mathbf{m}_1)(\mathbf{m} - \mathbf{m}_1)' + \frac{n_0 n}{n_1}(\bar{\mathbf{x}} - \mathbf{m}_0)(\bar{\mathbf{x}} - \mathbf{m}_0)'$$

Scanned with OKEN Scanner

și deci expresia (14.6.16) devine

$$\left\{(\det \delta^2)^{-k/2} \exp\left[-\frac{n_1}{2}(\mathrm{m}-\mathrm{m}_1)'\,\delta^{-2}(\mathrm{m}-\mathrm{m}_1)\right]\right\} \times$$

$$\times\left\{(\det \delta^2)^{-\frac{\nu_1+k+1}{2}} \exp\left[-\frac{1}{2}\,\mathrm{ur}\,(\delta^{-2}\,\psi_1)\right]\right\},$$

ceea ce arată că densitatea de probabilitate aposteriori a lui $(\widetilde{\mathrm{m}}, \widetilde{\delta}^2)$ este $f_{NWi}(\mathrm{m}, \delta^2 \mid \mathrm{m}_1, n_1, \psi_1, \nu_1)$ și teorema este demonstrată.

Teorema 14.6.6 (Cazul E). *Fie* $\widetilde{\mathrm{x}}_1, \ldots, \widetilde{\mathrm{x}}_n$ *o selecție de volum* n *din* $\widetilde{\mathrm{x}} \in \mathrm{N}_k(\mathrm{m}, \delta^2)$, *unde* m *și* δ^2 *sînt necunoscuți.*

Dacă se adoptă pentru $(\widetilde{\mathrm{m}}, \widetilde{\delta}^2)$ *densitatea de probabilitate apriori* $\mathrm{f_{NWi}}(\mathrm{m}, \delta^2 \mid \mathrm{m}_0, \mathrm{n}_0, \psi_0, \nu_0)$, *atunci:*

(a) Densitatea de probabilitate condiționată $\mathrm{f}(\overline{\mathrm{x}} \mid \mathrm{s}^2)$ *este dată de relația*

$$\mathrm{f}(\overline{\mathrm{x}} \mid \mathrm{s}^2) = f_S\left(\overline{\mathrm{x}} \mid \mathrm{m}_0, \frac{\mathrm{n}_1}{n_0(\nu_1-\mathrm{k})}\psi_2, \nu_1-\mathrm{k}\right), \quad (14.6.18)$$

unde

$$\psi_2 = \frac{1}{\mathrm{n}}\psi_0 + \frac{\mathrm{n}-1}{\mathrm{n}}\mathrm{s}^2; \quad (14.6.19)$$

(b) Densitatea de probabilitate $\mathrm{f}(\overline{\mathrm{x}})$ *este dată de relația*

$$\mathrm{f}(\overline{\mathrm{x}}) = \mathrm{f_S}\left(\overline{\mathrm{x}} \mid \mathrm{m}_0, \frac{\mathrm{n}_1}{\mathrm{n}_0\mathrm{n}(\nu_0-\mathrm{k}+1)}\psi_0, \nu_0-\mathrm{k}+1\right); \quad (14.6.20)$$

(c) Densitatea de probabilitate condiționată $\mathrm{f}(\mathrm{s}^2 \mid \overline{\mathrm{x}})$ *este dată de relația*

$$\mathrm{f}(\mathrm{s}^2 \mid \overline{\mathrm{x}}) = \mathrm{f}_{\beta 2i}(\mathrm{s}^2 \mid \psi_3, \nu_0+\mathrm{k}+2, \mathrm{n}-1, \mathrm{n}-1), \quad (14.6.21)$$

unde

$$\psi_3 = \psi_0 + \frac{\mathrm{n}_0\mathrm{n}}{\mathrm{n}_1}(\overline{\mathrm{x}}-\mathrm{m}_0)(\overline{\mathrm{x}}-\mathrm{m}_0)'; \quad (14.6.22)$$

Scanned with OKEN Scanner

(d) Densitatea de probabilitate $f(s^2)$ *este dată de relația*

$$f(s^2) = f_{\beta 2i}(s^2 \mid \psi_0,\ \nu_0,\ n-1,\ n-1). \qquad (14.6.23)$$

Demonstrație. Pentru m și δ^2 dați, avem $\tilde{\bar{x}} \in N_k\left(m, \frac{1}{n}\delta^2\right)$, $\tilde{s}^2 \in W((n-1)\delta^2, n-1)$, iar $\tilde{\bar{x}}$ și $\tilde{s}^2$ sînt independenți, conform teoremei 13.9.3. Urmează atunci că densitatea de probabilitate a lui $(\tilde{\bar{x}}, \tilde{s}^2)$ condiționată de m și δ^2 este de forma

$$f(\bar{x},\ s^2 \mid m,\ \delta^2) = f_N\left(\bar{x} \mid m,\ \frac{1}{n}\delta^2\right) f_W(s^2 \mid (n-1)\delta^2,\ n-1).$$

Dacă densitatea de probabilitate apriori a lui $(\tilde{m}, \tilde{\delta}^2)$ este $f_{NWi}(m, \delta^2 \mid m_0, n_0, \psi_0, \nu_0)$, atunci, conform teoremei precedente, rezultă că densitatea de probabilitate aposteriori a lui $(\tilde{m}, \tilde{\delta}^2)$ este $f_{NWi}(m, \delta^2 \mid m_1, n_1, \psi_1, \nu_1)$ și deci densitatea de probabilitate $f(\bar{x}, s^2)$ se poate scrie sub forma

$$f(\bar{x},\ s^2) = \frac{f(\bar{x},\ s^2 \mid m,\ \delta^2)\, f(m,\ \delta^2)}{f(m,\ \delta^2 \mid \bar{x},\ s^2)} =$$

$$= f_N\left(\bar{x} \mid m,\ \frac{1}{n}\delta^2\right) \frac{f_W(s^2 \mid (n-1)\delta^{-2}, n-1)\, f_{NWi}(m, \delta^2 \mid m_0, n_0, \psi_0, \nu_0)}{f_{NWi}(m,\ \delta^2 \mid m_1, n_1, \psi_1, \nu_1)} \qquad (14.6.24)$$

Vom demonstra mai întîi afirmația (b) a teoremei. Densitatea de probabilitate $f(\bar{x})$ se obține în modul următor:

$$f(\bar{x})) = \iint f(\bar{x} \mid m,\ \delta^2)\, f(m,\ \delta^2)\, dm\, d\delta^2 =$$

$$= \iint f_N\left(\bar{x} \mid m,\ \frac{1}{n}\delta^2\right) f_{NWi}(m,\ \delta^2 \mid m_0,\ n_0,\ \psi_0,\ \nu_0)\, dm\, d\delta^2 =$$

$$= \int \left[\int f_N\left(\bar{x} \mid m,\ \frac{1}{n}\delta^2\right) f_N\left(m \mid m_0,\ \frac{1}{n_0}\delta^2\right) dm\right] f_{Wi}(\delta^2 \mid \psi_0,\ \nu_0)\, d\delta^2.$$

Scanned with OKEN Scanner

Conform teoremei 14.6.1, integrala din paranteza dreaptă de mai sus este

$$f_N\left(\bar{\mathbf{x}} \mid \mathbf{m}_0,\ \frac{1}{n_0}\,\delta^2 + \frac{1}{n}\,\delta^2\right) = f_N\left(\bar{\mathbf{x}} \mid \mathbf{m}_0,\ \frac{n_1}{n_0 n}\,\delta^2\right)$$

și deci

$$f(\bar{\mathbf{x}}) = \int f_N\left(\bar{\mathbf{x}} \mid \mathbf{m}_0,\ \frac{n_1}{n_0\, n}\,\delta^2\right) f_{Wi}(\delta^2 \mid \psi_0,\ \nu_0)\, d\delta^2 =$$

$$= \int f_{NWi}\left(\bar{\mathbf{x}},\ \delta^2 \mid \mathbf{m}_0,\ \frac{n_0\, n}{n_1},\ \psi_0,\ \nu_0\right) d\delta^2.$$

Aplicînd acum teorema 14.6.4, rezultă că ultima integrală are valoarea

$$fs\left(\bar{\mathbf{x}} \mid \mathbf{m}_0,\ \frac{n_1}{n_0 n(\nu_0 - k + 1)}\,\psi_0,\ \nu_0 - k + 1\right),$$

adică tocmai ceea ce trebuia obținut.

Pentru demonstrația afirmației *(c)*, să observăm mai întîi că

$$f(\mathbf{s}^2 \mid \bar{\mathbf{x}}) = \frac{f(\bar{\mathbf{x}},\ \mathbf{s}^2)}{f(\bar{\mathbf{x}})},$$

de unde se obține ușor că $f(\mathbf{s}^2 \mid \bar{\mathbf{x}})$ este de forma

$$k_0(\det \mathbf{s}^2)^{\frac{n-k}{2}-1}\left[\det\left(\mathbf{s}^2 + \frac{1}{n-1}\psi_3\right)\right]^{\frac{\nu_1+k+1}{2}}, \quad (14.6.25)$$

unde k_0 este un factor independent de $\mathbf{s}^2$. Din ultima formă a lui $f(\mathbf{s}^2 \mid \bar{\mathbf{x}})$ se vede ușor că această densitate de probabilitate este $f_{\beta 2i}(\mathbf{s}^2 \mid \psi_3,\ \nu_0 + k + 2,\ \mathrm{n}-1,\ \mathrm{n}-1)$, adică tocmai ceea ce trebuia demonstrat.

Pentru a demonstra afirmația *(d)* a teoremei, să observăm mai întîi că pentru $\mathbf{m}$ și δ^2 dați, $\tilde{\mathbf{s}}^2$ are densitatea de probabilitate $f_W(\mathbf{s}^2 \mid (n-1)\,\delta^{-2},\ n-1)$. Atunci rezultă că

$$f(\mathbf{s}^2) = \iint f(\mathbf{s}^2 \mid \mathbf{m},\ \delta^2)\, f(\mathbf{m},\ \delta^2)\, d\mathbf{m}\, d\delta^2 =$$

$$= \iint f_W(s^2 \mid (n-1)\delta^{-2},\ n-1)\, f_{NWi}(m,\ \delta^2 \mid m_0,\ n_0,\ \psi_0,\ \nu_0)\, dm\, d\delta^2 =$$

$$= \iint f_W(s^2 \mid (n-1)\,\delta^{-2},\ n-1)\, f_N\left(m \mid m_0,\ \frac{1}{n_0}\delta^2\right) f_{Wi}(\delta^2 \mid \psi_0, \nu_0)\, dm d\delta^2 =$$

$$= \int f_W(s^2 \mid (n-1)\,\delta^{-2},\ n-1)\, f_{Wi}(\delta^2 \mid \psi_0,\ \nu_0)\, d\delta^2.$$

Utilizînd apoi teorema 13.8.1 rezultă că ultima integrală are valoarea $f_{\beta 2i}(s^2 \mid \psi_0,\ \nu_0,\ n-1,\ n-1)$ și afirmația *(d)* este demonstrată.

În sfîrșit, pentru a demonstra afirmația *(a)*, să observăm că

$$f(\bar{x} \mid s^2) = \frac{f(\bar{x},\ s^2)}{f(s^2)},$$

unde $f(\bar{x},\ s^2)$ este dat de relația (14.6.24), iar $f(s^2)$ este dat de relația (14.6.23). Rezultă atunci că $f(\bar{x}|s^2)$ este de forma

$$k_0'(\det \psi_1)^{-\frac{\nu_1}{2}}$$

unde ψ_1 este dat de (14.6.15), iar k_0' este un factor care nu depinde de $\bar{x}$. Dacă notăm

$$\psi_4 = \frac{n_1}{n_0\, n}\,[\psi_0 + (n-1)s^2],$$

rezultă că $f(\bar{x}|s^2)$ este de forma

$$k_0''[\det(\psi_4 + (\bar{x}-m_0)(\bar{x}-m_0)')]^{-\frac{\nu_1}{2}},$$

unde k_0'' este un factor care nu depinde de $\bar{x}$. Utilizînd din nou relația (14.6.11) pentru $A = \psi_4$ și $x = \bar{x}-m_0$, rezultă că

$$\det(\psi_4 + (\bar{x}-m_0)(\bar{x}-m_0)') = (\det \psi_4)\,[1 +$$

$$+ (\bar{x}-m_0)'\,\psi_4^{-1}(\bar{x}-m_0)],$$

ceea ce arată că $f(\bar{x} \mid s^2)$ este de forma

$$k_0''' [1 + (\bar{x} - m_0)' \psi_4^{-1} (\bar{x} - m_0)]^{-\frac{\nu_1}{2}},$$

unde k_0''' este un factor care nu depinde de $\bar{x}$. Rezultă atunci că

$$f(\bar{x} \mid s^2) = f_S\left(\bar{x} \mid m_0, \frac{1}{\nu_1 - k} \psi_4, \nu_1 - k\right).$$

Observînd însă că

$$\frac{1}{\nu_1 - k} \psi_4 = \frac{n_1}{n_0(\nu_1 - k)} \psi_2,$$

rezultatul căutat (14.6.18) se deduce imediat din ultimele două relații obținute mai sus.

În cazul cînd m și σ_*^2 sînt necunoscuți, iar δ_*^2 este cunoscut (Cazul B considerat în § 11.7.1), densitatea de probabilitate conjugată pentru $(\widetilde{m}, \widetilde{\sigma}_*^2)$ se dovedește a fi așa-numita densitate de probabilitate *normală-gama inversă*, notată prin $f_{N\gamma i}(m, \sigma_*^2 | m_0, \delta_{0*}^2, \psi_0, \nu_0)$ și definită ca fiind egală cu produsul

$$f_N(m \mid m_0, \sigma_*^2 \delta_{0*}^2)\, f_{\gamma i}(\sigma_*^2 \mid \psi_0, \nu_0). \qquad (14.6.26)$$

Teorema 14.6.7 (Cazul B). *Dacă* $(\widetilde{m}, \widetilde{\sigma}_*^2)$ *are densitatea de probabilitate* $f_{N\gamma i}(m, \sigma_*^2 \mid m_0, \delta_{0*}^2, \psi_0, \nu_0)$, *atunci densitatea de probabilitate (marginală) pentru* $\widetilde{m}$ *este* $f_S(m \mid m_0, \psi_0 \delta_{0*}^2, \nu_0)$, *iar densitatea de probabilitate (marginală) a lui* $\widetilde{\sigma}_*^2$ *este* $f_{\gamma i}(\sigma_*^2 \mid \psi_0, \nu_0)$.

Demonstrație. Deoarece $\widetilde{\sigma}_*^2 \in \gamma i(\psi_0, \nu_0)$ dacă și numai dacă variabila aleatoare $\widetilde{x} = 1/\widetilde{\sigma}_*^2$ aparține clasei $\gamma 2(\psi_0, \nu_0)$, rezultă că

$$f(m) = \int f_N(m | m_0, \sigma_*^2 \delta_{0*}^2)\, f_{\gamma i}(\sigma_*^2 \mid \psi_0, \nu_0)\, d\sigma_*^2 =$$

$$= \int f_N\left(m | m_0, \frac{1}{x} \delta_{0*}^2\right) f_{\gamma 2}(x \mid \psi_0, \nu_0)\, dx.$$

Conform teoremei 13.6.1, ultima integrală are valoarea

$$f_S(m \mid m_0, \psi_0, \delta_{0*}^2, \nu_0)$$

și deci prima afirmație a teoremei este demonstrată.

A doua afirmație a teoremei este evidentă.

Teorema 14.6.8 (Cazul B). *Să considerăm o selecție de volum* n *din* $\tilde{x} \in N_k(m, \sigma_*^2 \delta_*^2)$, *unde* m *și* σ_*^2 *sînt necunoscuți, iar* δ_*^2 *este cunoscut.*

Dacă densitatea de probabilitate apriori a lui $(\tilde{m}, \tilde{\sigma}_*^2)$ *este* $f_{N\gamma i}(m, \sigma_*^2 \mid m_0, \delta_{0*}^2, \psi_0, \nu_0)$, *atunci densitatea de probabilitate aposteriori a lui* $(\tilde{m}, \tilde{\sigma}_*^2)$ *este* $f_{N\gamma i}(m, \sigma_*^2 \mid m'_1, \delta_{1*}^2, \psi_1, \nu_1)$, *unde*

$$m_1 = \delta_{1*}^2(\delta_{0*}^{-2} m_0 + n \delta_*^{-2} \bar{x}), \tag{14.6.27}$$

$$\delta_{1*}^2 = (\delta_{0*}^{-2} + n \delta_*^{-2})^{-1}, \tag{14.6.28}$$

$$\psi_1 = \frac{1}{\nu_1}\Big[\psi_0 \nu_0 + k(n-1)s_*^2 + (\bar{x} - m_0)' \left(\delta_{0*}^2 + \frac{1}{n}\delta_*^2\right)^{-1} (\bar{x} - m_0)\Big] \tag{14.6.29}$$

și

$$\nu_1 = \nu_0 + kn. \tag{14.6.30}$$

Demonstrație. Conform teoremei 11.7.2, în acest caz statistica $(\tilde{n}, \tilde{\bar{x}}, \tilde{s}_*^2)$ este suficientă pentru (m, σ_*^2), iar funcția de verosimilitate are forma

$$L(m, \sigma_*^2 \mid n, \bar{x}, s_*^2) =$$

$$= k_0 (\sigma_*^2)^{-\frac{kn}{2}} \exp\left[-\frac{k(n-1)}{2\sigma_*^2} s_*^2\right] \exp\left[-\frac{n}{2\sigma_*^2}(\bar{x}-m)'\delta_*^{-2}(\bar{x}-m)\right],$$

unde

$$k_0 = (2\pi)^{-\frac{kn}{2}}.$$

Dacă densitatea de probabilitate apriori a lui $(\tilde{m}, \tilde{\sigma}_*^2$ este $f_{N\gamma i}(m, \sigma_*^2 \mid m_0, \delta_{0*}^2, \psi_0, \nu_0)$, atunci densitatea de probabilitate aposteriori pentru $(\tilde{m}, \tilde{\sigma}_*^2)$ este proporțională cu

$$f_N(m \mid m_0, \sigma_*^2 \delta_{0*}^2)\, f_{\gamma i}(\sigma_*^2 \mid \psi_0, \nu_0)\, L(m, \sigma_*^2 \mid n, \bar{x}, s_*^2),$$

adică este proporțională cu

$$\sigma_*^2)^{-\frac{kn+k+\nu_0+1}{2}} \exp\left\{-\frac{1}{2\sigma_*^2}[\psi_0 \nu_0 + k(n-1)s_*^2 + \psi^*]\right\}, \tag{14.6.31}$$

unde

$$\psi^* = (\mathrm{m}-\mathrm{m}_0)'\ \delta_{0*}^{-2}(\mathrm{m}-\mathrm{m}_0) + n(\bar{\mathrm{x}}-\mathrm{m})'\ \delta_*^{-2}(\bar{\mathrm{x}}-\mathrm{m}).$$

Un calcul simplu ne conduce la concluzia că

$$\psi^* = (\mathrm{m}-\mathrm{m}_1)'\ \delta_{1*}^{-2}(\mathrm{m}-\mathrm{m}_1) +$$

$$+ (\bar{\mathrm{x}}-\mathrm{m}_0)'\left(\delta_{0*}^2 + \frac{1}{n}\delta_*^2\right)^{-1}(\bar{\mathrm{x}} - \mathrm{m}_0),$$

ceea ce arată că (14.6.31) poate fi scris sub forma

$$(\sigma_*^2)^{-k/2}\exp\left[-\frac{1}{2\sigma_*^2}(\mathrm{m} - \mathrm{m}_1)'\ \delta_{1*}^{-2}(\mathrm{m} - \mathrm{m}_1)\right]\times$$

$$\times(\sigma_*^2)^{-\frac{\nu_1+1}{2}}\exp\left[-\frac{\psi_1\nu_1}{2\sigma_*^2}\right],$$

de unde rezultă că densitatea de probabilitate aposteriori a lui $(\widetilde{\mathrm{m}},\ \widetilde{\sigma}_*^2)$ este $f_{N\gamma i}(\mathrm{m},\ \sigma_*^2 \mid \mathrm{m}_1,\ \delta_{1*}^2,\ \psi_1,\ \nu_1)$ și teorema este demonstrată.

Teorema 14.6.9. (Cazul B). *Fie* $\widetilde{\mathrm{x}}_1, \ldots, \widetilde{\mathrm{x}}_n$ *o selecție de volum* n *din* $\widetilde{\mathrm{x}} \in \mathrm{N}_k(\mathrm{m},\ \sigma_*^2\ \delta_*^2)$ *unde* m *și* σ_*^2 *sînt necunoscuți, iar* δ_*^2 *este cunoscut.*

Dacă $(\widetilde{\mathrm{m}},\ \widetilde{\sigma}_*^2)$ *are densitatea de probabilitate apriori* $\mathrm{f}_{N\gamma i}(\mathrm{m},\ \sigma_*^2 \mid \mathrm{m}_0,\ \delta_0^2,\ \psi_0,\ \nu_0)$, *atunci:*

(a) Densitatea de probabilitate condiționată $\mathrm{f}(\bar{\mathrm{x}} \mid \mathrm{s}_*^2)$ *este dată de relația*

$$\mathrm{f}(\bar{\mathrm{x}} \mid \mathrm{s}_*^2) = \mathrm{f}_S(\bar{\mathrm{x}} \mid \mathrm{m}_0,\ \psi_*,\ \nu_0 + k(n-1)), \qquad (14.6.32)$$

unde

$$\psi_* = \left[\frac{\psi_0\,\nu_0 + k(n-1)s_*^2}{\nu_0 + k(n-1)}\right]\left(\delta_{0*}^2 + \frac{1}{n}\,\delta_*^2\right); \qquad (14.6.33)$$

(b) Densitatea de probabilitate $\mathrm{f}(\bar{\mathrm{x}})$ *este dată de relația*

$$\mathrm{f}(\bar{\mathrm{x}}) = \mathrm{f}_S\left(\bar{\mathrm{x}} \mid \mathrm{m}_0,\ \psi_0\left(\delta_{0*}^2 + \frac{1}{n}\,\delta_*^2\right),\ \nu_0\right); \qquad (14.6.34)$$

Scanned with OKEN Scanner

(c) Densitatea de probabilitate condiționată $f(s_*^2 \mid \bar{x})$ *este dată de relația*

$$f(s_*^2 \mid \bar{x}) = f_{\beta 2_i}\left(s_*^2 \,\middle|\, \frac{k(n-1)}{2},\ \frac{\nu_0 + k}{2} + 1,\ b\right), \quad (14.6.35)$$

unde

$$b = \frac{1}{k(n-1)}\left[\psi_0 \nu_0 + (\bar{x} - m_0)'\left(\delta_{0*}^2 + \frac{1}{n}\,\delta_*^2\right)^{-1}(\bar{x} - m_0)\right]; \quad (14.6.36)$$

(d) Densitatea de probabilitate $f(s_*^2)$ *este dată de relația*

$$f(s_*^2) = f_{\beta 2_i}\left(s_*^2 \,\middle|\, \frac{k(n-1)}{2},\ \frac{\nu_0}{2},\ \frac{\psi_0 \nu_0}{k(n-1)}\right). \quad (14.6.37)$$

Demonstrație. Conform teoremei 13.9.3, pentru m, σ_*^2 și δ_*^2 dați, statistica $\tilde{s}_*^2$ aparține clasei $\gamma 2(1/\sigma_*^2,\ k(n-1))$ și este independentă de $\tilde{\bar{x}}$, care aparține clasei $N\left(m, \frac{1}{n}\,\sigma_*^2 \delta_*^2\right)$. Rezultă atunci că densitatea de probabilitate a vectorului aleator $(\tilde{\bar{x}},\ \tilde{s}_*^2)$ condiționată de $(\tilde{m},\ \tilde{\sigma}_*^2)$ este

$$f(\bar{x},\ s_*^2 \mid m,\ \sigma_*^2) = f_N\left(\bar{x} \mid m,\ \frac{1}{n}\,\sigma_*^2 \delta_*^2\right) f_{\gamma 2}(s_*^2 \mid 1/\sigma_*^2,\ k(n-1)).$$

Dacă densitatea de probabilitate apriori a lui $(\tilde{m},\ \tilde{\sigma}_*^2)$ este $f_{N\gamma i}(m,\ \sigma_*^2 \mid m_0,\ \delta_{0*}^2,\ \psi_0,\ \nu_0)$, atunci, conform teoremei precedente, densitatea de probabilitate aposteriori a lui $(\tilde{m},\ \tilde{\sigma}_*^2)$ este $f_{N\gamma i}(m,\ \sigma_*^2 \mid m_1,\ \delta_{1*}^2,\ \psi_1,\ \nu_1)$ și deci densitatea de probabilitate $f(\bar{x},\ s_*^2)$ se poate scrie sub forma

$$f(\bar{x},\ s_*^2) = \frac{f(\bar{x},\ s_*^2 \mid m,\ \sigma_*^2)\ f(m,\ \sigma_*^2)}{f(m,\ \sigma_*^2 \mid \bar{x},\ s_*^2)} =$$

$$= \frac{f_N(\bar{x} \mid m,\ \frac{1}{n}\,\sigma_*^2 \delta_*^2)\, f_{\gamma 2}(s_*^2 \mid 1/\sigma_*^2,\ k(n-1))\ f_{N\gamma i}(m,\ \sigma_*^2 \mid m_0,\ \delta_{0*}^2,\ \psi_0,\ \nu_0)}{f_{N\gamma i}(m,\ \sigma_*^2 \mid m_1,\ \delta_{1*}^2,\ \psi_1,\ \nu_1)} \quad (14.6.38)$$

Ca și în cazul teoremei 14.6.6, vom demonstra mai întîi afirmația (*b*). Avem

$$f(\bar{x}) = \iint f(\bar{x} \mid m, \sigma_*^2)\, f(m, \sigma_*^2)\, dm\, d\sigma_*^2 =$$

$$= \iint f_N\left(\bar{x} \mid m, \frac{1}{n}\sigma_*^2\delta_*^2\right) f_{N\gamma i}(m, \sigma_*^2 \mid m_0, \delta_{0*}^2, \psi_0, \nu_0)\, dm\, d\sigma_*^2 =$$

$$= \int\left[\int f_N\left(\bar{x} \mid m, \frac{1}{n}\sigma_*^2\delta_*^2\right) f_N(m \mid m_0, \sigma_*^2\delta_{0*}^2)\, dm\right] f_{\gamma i}(\sigma_*^2 \mid \psi_0, \nu_0)\, d\sigma_*^2.$$

Conform teoremei 14.6.1, integrala din paranteza dreaptă de mai sus este egală cu $f_N\left(\bar{x} \mid m_0, \sigma_*^2\delta_{0*}^2 + \frac{1}{n}\sigma_*^2\delta_*^2\right)$ și deci putem scrie

$$f(\bar{x}) = \int f_N\left(\bar{x} \mid m_0, \sigma_*^2\left(\delta_{0*}^2 + \frac{1}{n}\delta_*^2\right)\right) f_{\gamma i}(\sigma_*^2 \mid \psi_0, \nu_0)\, d\sigma_*^2 =$$

$$= \int f_{N\gamma i}\left(\bar{x}, \sigma_*^2 \mid m_0, \delta_{0*}^2 + \frac{1}{n}\delta_*^2, \psi_0, \nu_0\right) d\sigma_*^2.$$

Conform teoremei 14.6.7, ultima integrală este egală cu

$$f_S\left(\bar{x} \mid m_0, \psi_0\left(\delta_{0*}^2 + \frac{1}{n}\delta_*^2\right), \nu_0\right),$$

ceea ce trebuia demonstrat.

Pentru a demonstra afirmația (*c*), să observăm că putem scrie

$$f(s_*^2 \mid \bar{x}) = \frac{f(\bar{x}, s_*^2)}{f(\bar{x})} =$$

$$= \frac{f_N\left(\bar{x} \mid m, \frac{1}{n}\sigma_*^2\delta_*^2\right) f_{\gamma 2}(s_*^2 \mid 1/\sigma_*^2, k(n-1))\, f_{N\gamma i}(m, \sigma_*^2 \mid m_0, \delta_{0*}^2, \psi_0, \nu_0)}{f_S\left(\bar{x} \mid m_0, \psi_0\left(\delta_{0*}^2 + \frac{1}{n}\delta_*^2\right), \nu_0\right) f_{N\gamma i}(m, \sigma_*^2 \mid m_1, \delta_{1*}^2, \psi_1, \nu_1)}$$

Scanned with OKEN Scanner

de unde se obține ușor că $f(s_*^2 \mid \overline{x})$ este de forma

$$f(s_*^2|\overline{x}) = k_0(s_*^2)^{\frac{k(n-1)}{2}-1}\left[\psi_0\,\nu_0 + k(n-1)s_*^2 + \right.$$

$$\left. + (\overline{x}-m_0)'\left(\delta_{0*}^2 + \frac{1}{n}\,\delta_*^2\right)^{-1}(\overline{x}-m_0)\right]^{-\frac{\nu_1+2}{2}},$$

unde k_0 este un factor care nu conține pe s_*^2. Rezultatul căutat se deduce ușor din ultima relație obținută mai sus.

Pentru a demonstra afirmația *(d)*, să observăm că, pentru m, σ_*^2 și δ_*^2 dați, densitatea de probabilitate a lui $\widetilde{s_*^2}$ este $f_{\gamma 2}(s_*^2 \mid 1/\sigma_*^2, k(n-1))$ și deci

$$f(s_*^2) = \iint f(s_*^2 \mid m, \sigma_*^2, \delta_*^2)\, f(m, \sigma_*^2)\, dm\, d\sigma_*^2 =$$

$$= \iint f_{\gamma 2}(s_*^2 \mid 1/\sigma_*^2, k(n-1))\, f_{N\gamma i}(m, \sigma_*^2 \mid m_0, \delta_{0*}^2, \psi_0, \nu_0)\, dm\, d\sigma_*^2 =$$

$$= \iint f_{\gamma 2}(s_*^2 \mid 1/\sigma_*^2, k(n-1))\, f_N(m \mid m_0, \sigma_*^2\delta_{0*}^2)\, f_{\gamma i}(\sigma_*^2 \mid \psi_0, \nu_0)\, dm\, d\sigma_*^2 =$$

$$= \int f_{\gamma 2}(s_*^2 \mid 1/\sigma_*^2, k(n-1))\, f_{\gamma i}(\sigma_*^2 \mid \psi_0, \nu_0)\, d\sigma_*^2.$$

După calcule elementare se obține

$$f(s_*^2) = f_{\beta 2 i}\left(s_*^2 \,\middle|\, \frac{k(n-1)}{2},\ \frac{\nu_0}{2},\ \frac{\psi_0\,\nu_0}{k(n-1)}\right),$$

ceea ce trebuia demonstrat.

În sfîrșit, pentru a demonstra afirmația *(a)* să observăm că, utilizînd relațiile (14.6.37) și (14.6.38), putem scrie

$$f(\overline{x} \mid s_*^2) = \frac{f(\overline{x}, s_*^2)}{f(s_*^2)} =$$

$$= \frac{f_N\left(\overline{x} \mid m, \frac{1}{n}\,\sigma_*^2\delta_*^2\right) f_{\gamma 2}(s_*^2 \mid 1/\sigma_*^2, k(n-1))\, f_{N\gamma i}(m, \sigma_*^2 \mid m_0, \delta_{0*}^2, \psi_0, \nu_0)}{f_{N\gamma i}(m, \sigma_*^2 \mid m_1, \delta_{1*}^2, \psi_1, \nu_1)\, f_{\beta 2 i}\left(s_*^2 \,\middle|\, \frac{k(n-1)}{2},\ \frac{\nu_0}{2},\ \frac{\psi_0\,\nu_0}{k(n-1)}\right)}.$$

După calcule elementare se obține relația (14.6.32) și teorema este astfel complet demonstrată.

ANEXĂ

A.1. FUNCȚIA GAMA

Funcția $\Gamma(r)$ a lui Euler este definită pentru orice $r > 0$ prin relația

$$\Gamma(r) = \int_0^\infty x^{r-1} e^{-x} \, dx. \tag{A.1.1}$$

Făcînd o integrare prin părți se obține relația funcțională

$$\Gamma(r+1) = r \, \Gamma(r). \tag{A.1.2}$$

valabilă pentru orice $r > 0$. În particular, pentru r întreg, obținem

$$\Gamma(r+1) = r! \, \Gamma(1). \tag{A.1.3}$$

Observînd că

$$\Gamma(1) = \int_0^\infty e^{-x} \, dx = 1,$$

rezultă că, pentru r întreg, avem

$$\Gamma(r+1) = r! \tag{A.1.4}$$

Să observăm de asemenea că, utilizînd schimbarea de variabilă $2x = y^2$, obținem

$$\int_0^\infty y^{2r-1} e^{-y^2/2} \, dy = 2^{r-1} \, \Gamma(r) \tag{A.1.5}$$

Scanned with OKEN Scanner

și, în particular, pentru $r = 1/2$, obținem

$$\int_0^\infty e^{-y^2/2}\,dy = \frac{1}{\sqrt{2}}\,\Gamma\left(\frac{1}{2}\right).$$

Ținînd seama că

$$\int_0^\infty e^{-y^2/2}\,dy = \frac{1}{2}\int_{-\infty}^{\infty} e^{-y^2/2}\,dy = \frac{1}{2}\sqrt{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}\,dx\,dy} =$$

$$= \frac{1}{2}\sqrt{2\pi} = \sqrt{\pi/2},$$

rezultă că

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}. \tag{A.1.6}$$

Este acum clar că putem calcula ușor $\Gamma\left(\frac{2r+1}{2}\right)$ pentru r întreg, deoarece

$$\Gamma\left(\frac{2r+1}{2}\right) = \Gamma\left(r + \frac{1}{2}\right) = \Gamma\left(r - \frac{1}{2} + 1\right)$$

și deci

$$\Gamma\left(\frac{2r+1}{2}\right) = \left(r - \frac{1}{2}\right)\Gamma\left(r - \frac{1}{2}\right). \tag{A.1.7}$$

De exemplu, pentru $r = 1$, obținem

$$\Gamma\left(\frac{3}{2}\right) = \left(1 - \frac{1}{2}\right)\Gamma\left(1 - \frac{1}{2}\right) =$$

$$= \frac{1}{2}\,\Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2}. \tag{A.1.8}$$

A.2. FUNCȚIA BETA

Funcția beta a lui Euler $B(\rho, \nu)$ este definită pentru orice $\rho > 0$ și $\nu > 0$ prin relația

$$B(\rho, \nu) = \int_0^1 x^{\rho-1}(1-x)^{\nu-1}\,dx. \tag{A.2.1}$$

Făcînd schimbarea de variabilă $y = 1-x$, se observă ușor că

$$B(\rho, \nu) = B(\nu, \rho). \tag{A.2.2}$$

Să observăm de asemenea că substituția $x = \sin^2\theta$ în (A.2.1) ne conduce la relația

$$B(\rho, \nu) = 2\int_0^{\pi/2} \sin^{2\rho-1}\theta\, \cos^{2\nu-1}\theta\, d\theta, \tag{A.2.3}$$

iar substituția $x = 1/(1+y)$ ne conduce la relația

$$B(\rho, \nu) = \int_0^\infty \frac{y^{\nu-1}}{(1+y)^{\rho+\nu}}\,dy. \tag{A.2.4}$$

Relația dintre funcțiile beta și gama se obține cu ajutorul formulelor (A.1.5) și (A.2.3) în modul următor:

$$2^{\rho-1}\Gamma(\rho)\cdot 2^{\nu-1}\Gamma(\nu) = 2^{\rho+\nu-2}\int_0^\infty\int_0^\infty e^{-(x^2+y^2)/2}\, x^{2\rho-1}\, y^{2\nu-1}\,dx\,dy =$$

$$= 2^{\rho+\nu-2}\int_0^\infty e^{-r^2/2}\, r^{\rho+\nu+1}\,dr \int_0^{\pi/2} \sin^{2\rho-1}\theta\, \sin^{2\nu-1}\theta\,d\theta =$$

$$= 2^{\rho+\nu-2}\,\Gamma(\rho+\nu)\,B(\rho, \nu)$$

Scanned with OKEN Scanner

și deci

$$B(\rho, \nu) = \frac{\Gamma(\rho)\,\Gamma(\nu)}{\Gamma(\rho+\nu)}, \tag{A.2.5}$$

care este tocmai relația căutată.

Utilizînd (A.1.2) și (A.2.5) se pot deduce ușor relațiile funcționale

$$B(\rho+1, \nu) = \frac{\rho}{\rho+\nu} B(\rho, \nu), \tag{A.2.6}$$

$$B(\rho, \nu+1) = \frac{\nu}{\rho+\nu} B(\rho, \nu) \tag{A.2.7}$$

și deci

$$B(\rho+1, \nu) + B(\rho, \nu+1) = B(\rho, \nu). \tag{A.2.8}$$

În cazul particular cînd ρ și ν sînt întregi, din (A.1.4) și (A.2.5) se deduce ușor că

$$B(\rho+1, \nu+1) = \frac{\rho!\,\nu!}{(\rho+\nu+1)!}. \tag{A.2.9}$$

O formulă utilă pentru funcția gama este

$$\Gamma\left(r+\frac{1}{2}\right) = \frac{\sqrt{\pi}\,\Gamma(2r)}{2^{2r-1}\,\Gamma(r)}, \tag{A.2.10}$$

numită și *formula lui Legendre*. Această formulă se obține simplu observînd că

$$B(r, r) = \int_0^1 \left(\frac{1}{4} - \left(\frac{1}{2} - x\right)^2\right)^{r-1} dx =$$

$$= 2\int_0^{1/2} \left(\frac{1}{4} - \left(\frac{1}{2} - x\right)^2\right)^{r-1} dx,$$

care, în urma schimbării de variabilă $x = (1 - \sqrt{t})/2$, devine

$$B(r, r) = \frac{1}{2^{2r-1}} \int_0^1 t^{-\frac{1}{2}} (1-t)^{r-1} \, dt = \frac{1}{2^{2r-1}} B\left(\frac{1}{2}, r\right) =$$

$$= \frac{\Gamma\left(\frac{1}{2}\right)\Gamma(r)}{2^{2r-1}\,\Gamma\left(r + \frac{1}{2}\right)}$$

sau încă

$$B(r, r) = \frac{\sqrt{\pi}\,\Gamma(r)}{2^{2r-1}\,\Gamma\left(r + \frac{1}{2}\right)} \qquad \text{(A.2.11)}$$

și deci formula căutată (A.2.10) se obține imediat din (A.2.11) și (A.2.5).

A.3. FUNCȚIA LUI DIRICHLET

Funcția beta a lui Euler poate fi generalizată în modul următor. Fie $D \subset R^k$ domeniul definit astfel

$$D = \left\{ \mathbf{x} \in R^k : x_i > 0, \quad 1 \leqslant i \leqslant k, \sum_{i=1}^{k} x_i < 1 \right\}. \qquad \text{(A.3.1)}$$

Funcţia lui Dirichlet sau funcția beta k—dimensională $B_k(\boldsymbol{\rho}, \nu)$ unde $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_k)' \in R^k$ cu $\rho_i > 0$, $1 \leqslant i \leqslant k$, iar $\nu > 0$ este definită prin

$$B_k(\boldsymbol{\rho}, \nu) = \int_D x_1^{\rho_1 - 1} \ldots x_k^{\rho_k - 1} (1 - x_1 - \ldots - x_k)^{\nu - 1} \, dx_1 \ldots dx_k. \qquad \text{(A.3.2)}$$

Să efectuăm schimbarea de variabile

$$\begin{aligned} x_1 &= y_1, \\ x_2 &= y_2(1 - x_1), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ x_k &= y_k(1 - x_1 - \ldots - x_{k-1}) \end{aligned} \qquad \text{(A.3.3)}$$

pentru care jacobianul este

$$\frac{D(x_1, \ldots, x_k)}{D(y_1, \ldots, y_k)} = (1-y_1)^{k-1}(1-y_2)^{k-2} \ldots (1-y_{k-1}). \quad (A.3.4)$$

Deoarece

$$\begin{aligned} x_1 &= y_1, \\ x_2 &= y_2(1-y_1), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ x_k &= y_k(1-y_1) \ldots (1-y_{k-1}), \end{aligned} \quad (A.3.5)$$

și

$$1 - x_1 - \ldots - x_k = (1-y_1) \ldots (1-y_k),$$

rezultă că integrala (A.3.2) devine

$$B_k(\boldsymbol{\rho}, \nu) = \int_0^1 y_1^{\rho_1-1}(1-y_1)^{\rho_2+\ldots+\rho_k+\nu-1}\,dy_1 \cdot$$

$$\cdot \int_0^1 y_2^{\rho_2-1}(1-y_2)^{\rho_3+\ldots+\rho_k+\nu-1}\,dy_2 \ldots \int_0^1 y_k^{\rho_k-1}(1-y_k)^{\nu-1}\,dy_k =$$

$$= B(\rho_1, \rho_2+\ldots+\rho_k+\nu)\, B(\rho_2, \rho_3+\ldots+\rho_k+\nu) \ldots B(\rho_k, \nu).$$

Apoi, utilizînd (A.2.5), obținem

$$B_k(\boldsymbol{\rho}, \nu) = \frac{\Gamma(\rho_1)\,\Gamma(\rho_2+\ldots+\rho_k+\nu)}{\Gamma(\rho_1+\rho_2+\ldots+\rho_k+\nu)} \cdot \frac{\Gamma(\rho_2)\,\Gamma(\rho_3+\ldots+\rho_k+\nu)}{\Gamma(\rho_2+\ldots+\rho_k+\nu)} \ldots$$

$$\ldots \frac{\Gamma(\rho_k)\,\Gamma(\nu)}{\Gamma(\rho_k+\nu)},$$

de unde se deduce

$$B_k(\boldsymbol{\rho}, \nu) = \frac{\Gamma(\rho_1)\,\Gamma(\rho_2) \ldots \Gamma(\rho_k)\,\Gamma(\nu)}{\Gamma(\rho_1+\rho_2+\ldots+\rho_k+\nu)}, \quad (A.3.6)$$

formulă care generalizează relația (A.2.5).

Scanned with OKEN Scanner

Se poate observa ușor că relații de tipul (A.2.6), (A.2.7) și (A.2.8) se pot scrie și pentru funcția lui Dirichlet.

În sfîrșit, dacă $\rho_1, \ldots, \rho_k$ și ν sînt întregi, atunci (A.3.6) se poate scrie sub forma

$$B_k(\rho_1+1, \ldots, \rho_k+1, \nu+1) = \frac{\rho_1! \ldots \rho_k! \, \nu!}{(\rho_1 + \ldots + \rho_k + \nu)!}. \quad \text{(A.3.7)}$$

A.4. URMA UNEI MATRICE

Pentru orice matrice pătrată $\mathbf{A} = (a_{ij})_{1 \leqslant i, j \leqslant k}$ se definește *urma* acestei matrice și se notează cu $\mathrm{ur}(\mathbf{A})$. Prin definiție, $\mathrm{ur}(\mathbf{A})$ este suma elementelor situate pe diagonala principală a matricei $\mathbf{A}$, adică

$$\mathrm{ur}(\mathbf{A}) = \sum_{i=1}^{k} a_{ii}. \quad \text{(A.4.1)}$$

Dacă $\mathbf{A} = (a_{ij})_{1 \leqslant i, j \leqslant k}$ și $\mathbf{B} = (b_{ij})_{1 \leqslant i, j \leqslant k}$ sînt matrice de ordinul k, atunci este clar că

$$\mathrm{ur}(\mathbf{AB}) = \sum_{i,j=1}^{k} a_{ij} b_{ji}. \quad \text{(A.4.2)}$$

Se poate de asemenea vedea că, dacă $\mathbf{A}$, $\mathbf{B}$ și $\mathbf{C}$ sînt matrice de ordinul k, $\mathbf{x} \in R^k$, iar $a \in \mathbf{R}$, atunci

$$\mathrm{ur}(\mathbf{AB}) = \mathrm{ur}(\mathbf{BA}), \quad \text{(A.4.3)}$$

$$\mathrm{ur}(a\mathbf{AB}) = \mathrm{ur}(\mathbf{A} \cdot a\mathbf{B}) = a \cdot \mathrm{ur}(\mathbf{AB}), \quad \text{(A.4.4)}$$

$$\mathrm{ur}(\mathbf{A}(\mathbf{B}+\mathbf{C})) = \mathrm{ur}(\mathbf{AB}) + \mathrm{ur}(\mathbf{AC}), \quad \text{(A.4.5)}$$

și

$$\mathrm{ur}(\mathbf{Axx}') = \mathbf{x}'\mathbf{Ax}. \quad \text{(A.4.6)}$$

Scanned with OKEN Scanner

BIBLIOGRAFIE

1. AITCHISON, J., *Choice against chance: An Introduction to Statistical Decision Theory*, Addison-Wesley, Reading Massachusets, 1970.
2. ANDERSON, T. W., *An Introduction to Multivariate Statistical Analysis*, Wiley, New York, 1958.
3. BLACKWELL, D., GIRSHICK, M. A., *Theory of Games and Statistical Decisions*, Wiley, New York, 1954.
4. CHERNOFF, H., MOSES, L. E., *Elementary Decision Theory*, Wiley, New York, 1959.
5. CUCULESCU, I., *Curs de teoria probabilităților*, Tipografia Universității București, 1976 (litografiat).
6. EVANS, I. G., *Bayesian Estimation of Parameters of a Multivariate Normal Distribution*, Journal of the Royal Statistical Society, Series B, 27 (1965), pp. 279–283.
7. FELLER, W., *An Introduction to Probability Theory and Its Applications*, vol. 2, Wiley, New York, 1966.
8. FELLNER, W., *Probability and Profit*, Homewood, Ill., Richard D. Irwin, Inc., 1965.
9. FERGUSON, T. S., *Mathematical Statistics: a Decision-Theoretic Approach*, Academic Press, New York, 1967.
10. FISHBURN, P. C., *Decision and Value Theory*, Wiley, New York, 1964.
11. GNEDENKO, B. V., *The Theory of Probability*, Mir Publishers, Moskow, 1969.
12. GHIHMAN, I. I., SKOROHOD, A. V., *Teoria slucealnîh proțessov*, tom I, Izd. Nauka. Moskva, 1971.
13. HALMOS, P. R., *Measure Theory*, Van Nostrand, Princeton, New Jersey, 1950.
14. HARDY, G. H., *Divergent Series*, Oxford University Press, Oxford, 1949.
15. IOSIFESCU, M., MIHOC, GH., THEODORESCU, R., *Teoria probabilităților și statistica matematică*, Editura Tehnică, București, 1966.
16. JEFFREYS, H., *Theory of probability*, Clarendon Press, New York, 1961.
17. KAUFMANN, A., DESBAZEILLE, G., *Metoda drumului critic*, Editura Tehnică, București, 1972.

Scanned with OKEN Scanner

CUPRINS

Partea întîi 5
INTRODUCERE ELEMENTARĂ ÎN TEORIA DECIZIEI 5

Capitolul 1. FORMULAREA PROBLEMEI 7
1.1 Introducere 7
1.2 Exemple 13
1.3 Structura comună 20
1.4 Observații generale privind rezolvarea unei probleme de decizie 24
1.5 Scurt istoric 25
1.6 Pro și contra acestei teorii 28
1.7 Necesitatea unei meta-teorii 32

Capitolul 2. UN EXEMPLU 34
2.1 Introducere 34
2.2 Analiza preliminară a problemei 36
2.3 Considerarea unei repartiții apriori 38
2.4 Clarviziune 42
2.5 Experimentare 44

Partea a doua 49
INCERTITUDINE 49

Capitolul 3. CÎMPURI DE PROBABILITATE 51
3.1 Introducere 51
3.2 Evenimente. Corpuri de evenimente 52
3.3 Definiția clasică a probabilității 58
3.4 Definiția axiomatică a probabilității 61
3.5. Probabilitate condiționată 69
3.6 Evenimente dependente și evenimente independente 77

Capitolul 4. VARIABILE ALEATOARE 81
4.1 Definiții și proprietăți 81
4.2 Funcția de repartiție a unei variabile aleatoare 88
4.3 Vectori aleatori. Funcții de repartiție multidimensionale 103
4.4 Variabile aleatoare independente 109

Scanned with OKEN Scanner

Capitolul 5. VALORI NUMERICE ASOCIATE VARIABILELOR ALEATOARE 114
5.1 Integrala Lebesgue 114
5.2 Media unei variabile aleatoare 123
5.3 Dispersia unei variabile aleatoare 130
5.4 Momente. Momente centrate 133
5.5 Mediană, cuantile, mod, asimetrie și exces 139
5.6 Corelație. Coeficient de corelație 141
5.7 Valori numerice asociate unui vector aleator 143

Capitolul 6. VALORI MEDII CONDIȚIONATE 146
6.1 Cazuri particulare 146
6.2 Valoare medie condiționată de un corp borelian 151
6.3 Valoare medie condiționată de o variabilă aleatoare 156
6.4 Funcții de repartiție condiționate 159
6.5 Momente condiționate 164
6.6 Funcții de regresie 166

Capitolul 7. ȘIRURI DE VARIABILE ALEATOARE 173
7.1 Convergența aproape sigură 173
7.2 Convergența în probabilitate 175
7.3 Convergența în medie 180
7.4 Convergența în repartiție 181
7.5 Compararea tipurilor de convergență 190

Capitolul 8. FUNCȚII CARACTERISTICE 191
8.1 Introducere 191
8.2 Proprietăți ale funcției caracteristice 193
8.3 Formula de inversiune 196
8.4 Legătura dintre funcția caracteristică și produsul de compoziție 201
8.5 Teoreme de convergență 203
8.6 Funcții generatoare 208

Capitolul 9. LEGEA NUMERELOR MARI 220
9.1 Introducere 220
9.2 Formularea problemei 221
9.3 Legea slabă a numerelor mari 222
9.4 Legea tare a numerelor mari 229

Capitolul 10. REPARTIȚIA NORMALĂ ȘI TEOREMA LIMITĂ CENTRALĂ 239
10.1 Introducere 239

Scanned with OKEN Scanner

10.2 Repartiția normală unidimensională 240
10.3 Repartiția normală multidimensională 245
10.4 Problema asimptotică centrală 262
10.5 Teorema limită centrală 263
10.6 Teoreme limită locale 274

Capitolul 11. INTRODUCERE ÎN STATISTICA MATEMATICĂ 279
11.1 Introducere 279
11.2 Selecția aleatoare dintr-o populație 279
11.3 Procese generatoare de date și procese de oprire 282
11.4 Inferență statistică. Statistici suficiente 285
11.5 Statistici uzuale 294
11.6 Procesul normal unidimensional 299
11.7 Procesul normal multidimensional 304

Capitolul 12. REPARTIȚII CLASICE DISCRETE 311
12.1 Introducere 311
12.2 Repartiția degenerată 311
12.3 Procesul Bernoulli și repartiții asociate 313
12.4 Procesul Bernoulli multidimensional 324
12.5 Procesul hipergeometric 331
12.6 Procesul hipergeometric multidimensional 339
12.7 Procesul Poisson 345

Capitolul 13. REPARTIȚII CLASICE CONTINUE 354
13.1 Introducere 354
13.2 Familia de repartiții gama 354
13.3 Familia de repartiții beta 364
13.4 Repartiția Student 375
13.5 Repartiția beta multidimensională 384
13.6 Repartiția Student multidimensională 391
13.7 Repartiția Wishart 395
13.8 Repartiția beta-2 inversă multidimensională 401
13.9 Repartițiile statisticilor asociate unui proces normal de generare de date 404

Capitolul 14. FAMILII DE REPARTIȚII CONJUGATE 420
14.1 Introducere 420
14.2 Procesul Bernoulli 421
14.3 Procesul Bernoulli multidimensional 426

Scanned with OKEN Scanner

14.4 Procesul Poisson 429
14.5 Procesul normal unidimensional 433
14.6 Procesul normal multidimensional 441

ANEXĂ

A.1 Funcția gama 460
A.2 Funcția beta 462
A.3 Funcția lui Dirichlet 464
A.4 Urma unei matrice 466

BIBLIOGRAFIE 467

Redactor: MARIA BORICEAN
Tehnoredactor: CONSTANTIN IORDACHE
Coli de tipar: 29,5. Bun de tipar: 26.06.1980.

Tiparul executat sub comanda nr. 1/133 la
Intreprinderea poligrafică
„13 Decembrie 1918",
str. Grigore Alexandrescu nr. 89—97
București,
Republica Socialistă România

Scanned with OKEN Scanner

Aspectele delicate și dificile ale unei probleme de decizie sînt legate de incertitudinea contextului și preferințele decidentului privind consecințele actului decizional. Necesitatea codificării informației privind incertitudinea într-o formă potrivită în scopul luării deciziilor a dus la dezvoltarea unor capitole importante ale teoriei probabilităților și statisticii matematice, precum și la schimbarea viziunii asupra posibilităților de aplicare în practică a acestor teorii.

Codificarea preferințelor decidentului privind consecințele actului decizional a dus la elaborarea *teoriei utilității*, care a înregistrat recent succese remarcabile.

Deși multe puncte ale teoriei deciziei sînt incomplet elaborate, iar tehnologia de aplicare la situații reale este încă destul de săracă în realizări importante, interesul crescînd pentru aceste metode stă la baza eforturilor actuale îndreptate spre rezolvarea acestor probleme.

Scanned with OKEN Scanner

Această lucrare în două volume prezintă teoria matematică a luării deciziilor în condiții de incertitudine. Primul volum este consacrat în esență problemelor de codificare a informației privind incertitudinea care poate să apară într-o problemă de decizie.

După o introducere privind problematica teoriei deciziei, sînt prezentate elementele de calculul probabilităților necesare în acest context, insistîndu-se în mod deosebit asupra valorilor medii condiționate. Expunerea continuă cu o introducere în statistica matematică, într-o viziune bayesiană, pentru a face apoi posibilă analiza principalelor repartiții clasice, accentuîndu-se asupra noțiunii de proces generator de date. Ultimul capitol prezintă pe scurt familiile importante de repartiții conjugate, utile în rezolvarea problemelor de inferență bayesiană.

Scanned with OKEN Scanner